南京大学大理科丛书

近代应用数学基础

苏维宜 著

清华大学出版社

南京大学大理科丛书

近代应用数学基础

苏维宜　著

清华大学出版社
北　京

内 容 简 介

本书结合例题，系统地介绍集合论、近世代数、点集拓扑、泛函分析、分布理论、微分几何等近代应用数学的基本内容及其在自然科学研究中的应用。书中强调对近代数学概念的理解和对重要论证方法的思路分析，以帮助读者掌握数学推理的基本思维方法，学会把近代应用数学中的重要定理和方法应用到本专业的具体问题中去。

本书可作为物理、天文、化学、地学、生物、计算机等专业学习相关课程的教材或参考书，也可供相关领域科研人员阅读参考。

图书在版编目(CIP)数据

近代应用数学基础/苏维宜著. --北京：清华大学出版社，2015
(南京大学大理科丛书)
ISBN 978-7-302-33820-8

Ⅰ. ①近…　Ⅱ. ①苏…　Ⅲ. ①应用数学　Ⅳ. ①O29

中国版本图书馆 CIP 数据核字(2013)第 212252 号

责任编辑：石　磊　陈　明
封面设计：
责任校对：刘玉霞
责任印制：

出版发行：清华大学出版社
网　　址：http://www.tup.com.cn，http://www.wqbook.com
地　　址：北京清华大学学研大厦 A 座　　**邮　　编：**100084
社 总 机：010-62770175　　**邮　　购：**010-62786544
投稿与读者服务：010-62776969，c-service@tup.tsinghua.edu.cn
质量反馈：010-62772015，zhiliang@tup.tsinghua.edu.cn
印 刷 者：
装 订 者：
经　　销：全国新华书店
开　　本：185mm×260mm　　**印　　张：**22　　**字　　数：**489 千字
版　　次：2015 年 4 月第 1 版　　**印　　次：**2015 年 4 月第1 次印刷
印　　数：1～　000
定　　价：　.00 元

产品编号：050950-01

南京大学大理科丛书

总　序

南京大学从1989年起实施大理科培养模式。实施的平台从基础学科教学强化部、基础学科教育学院到学院命名为匡亚明学院，规模不断扩大。在这种模式下，主修数学和理科不同一级学科（物理、化学、天文、地质地理、生物）的学生混合组班，构成了特殊的氛围。我们强调以学生为中心，实行学生自主选择学科方向的举措，“多次选择、逐步到位”，为学生的个性化发展提供了基本条件。学生比在一般按学科设置的院系中有更多的选择。学生可以自行决定是否选择学科交叉以及交叉的程度等。有相当数量的学生**自行设计成才之路**，这是一种近乎理想的境界。

高等学校理科人才培养历来采取“按学科设系”的格局，这几乎成为一种传统。20世纪50年代以后又出现了“专业”、“专门化”。这两种做法与自然科学本身发展的趋势有密切的关系。随着人们对自然的认识逐步深入，各学科及分支学科逐步建立，在每个学科领域又层层深入、越分越细。还原论的观点是发展过程的基本指导原则。然而在解释自然现象的过程中却时不时需要某种“综合”、多个学科或学科分支的配合，毕竟自然现象错综复杂，并不是按照人们的意志分门别类地发生。逐渐人们开始关注学科间的交叉、融合、综合以及自下层向上层的**演生**现象并形成相应理论。关于演生论请看物理学家对于物理学领域的评论[①]：

“在物理学过去的发展历史中，还原论的观点一直是物理学工作者进行研究的最基本的指导原则。它对整个学科的发展起到了巨大的推动作用，并取得了辉煌的成就。但是，以还原论为基础来研究和讨论复杂系统的合作现象时，却遇到了前所未有的挑战，从而使演生论的思想孕育而生，并成为当今物理学研究的重要指导原则。”

而学科交叉早在20世纪50年代就取得了令人瞩目的成绩，例如，华生和克里克（J. D. Watson，F. H. Crick）的合作导致了20世纪生命科学的最伟大的成就：DNA的双螺旋结构。随着时代的进步，学科交叉出现了两个明显的特征：**大规模团队作业和个人知识结构的丰富**。1997年物理学诺贝尔奖得主朱棣文领导了Biox这样的项目，而在第4届世界华人物理学家大会（Oversea Chinese Physicists Association 4）上他作的报告是“What can physics say about life?”这似乎彰显个人知识出现学科交叉的魅力。为这样的人才或对人

---

① 张广铭，于渌．物理学中的演生现象[J]．物理．2010，39(8)：543-549．

才的这种知识结构的需求提供机会和环境是顺应时代的要求。大理科模式的20年实践初步证明了这种选择是合时宜的。

本套丛书涵盖了大理科培养平台上所用的自行撰写的教材,包括数学和相关的自然科学。其中既有用于**拓宽**类的课程,如非生物学科的大学生物学,数理类的大学化学,化学生物类的理论物理、数理方法,数理类非天文学科的大学天文学,以及化学生物类和数理类的物理实验、化学实验等;也有用于**提升**类的课程:把传统的外系课程提升到交叉学科课程。教材既能满足传统的单纯学科方向的需要,又可供交叉选课采用。尽管当前传统型人才培养还是主流,但越来越多的人已经开始关注学科交叉的知识结构。希望本套丛书能够为有志于从事学科交叉的尝试提供参考。

教材的学术性、先进性是我们重点关注的。除了提供物质基础让学生通过选课构建交叉的知识结构外,丛书的作者也在教材内注意相关学科、相邻学科的链接,也在不同程度上注意物质世界层次间还原和演生的关系。至于学术层面深度融合的方面虽然出现了一些成果,但在教育层面还无法完全同步,需要时间研究、酝酿和采纳。

本套丛书由清华大学出版社出版发行。各书均经教学实践检验,先成先出,相对独立,文责自负。

卢德馨

2014年4月

前　　言

21 世纪飞速发展的科学技术与人类的生产实践，对科技人员的素质、知识与能力提出了新的要求。自然科学、社会科学等各个领域中的从业者，在数学思维水平、数学科学知识、数学应用能力方面所具备的基础，也达到了前所未有的高度。高等数学远远不能满足新的需求，近代数学的思维、概念、理论、方法已经悄然渗透，由高端变为基础，由理论变为现实，由指导性的方向变为科学中的实践。于是，继高等数学之后，一本介绍近代应用数学的教程编写迫在眉睫。

近代数学所包含的范围之广，知识面之宽，内容之深，非简单几句话所能概括。为了给自然科学的主体类（如物理学、天文学、化学、计算机科学、地球科学、生命科学等）的学生准备必要的近代数学知识，南京大学在 20 世纪 90 年代，首先在基础学科教学强化部开设了继高等数学之后的半年近代数学的课程，但未正式命名，而是作为高等数学的第四个学期而设置的，周学时为 5 的必修课。其内容涵盖勒贝格积分、微分几何等，使学生受益匪浅。

随着教学改革的深入和新世纪的到来，我们把勒贝格积分编写到高等数学教程中，把非数学类学生所需的近代数学知识集中在一起，从 2006 年开始编写教材，2007 年、2009 年两次印制成讲义，并在南京大学匡亚明学院开设了为时半年的近代应用数学课程，这就是本书的雏形。对于匡亚明学院的理科强化部、物理、天文类的学生，我们采用边教边修改教学内容的方式，逐步完善而成为目前的书稿。

本书的内容安排如下：

第 1 章介绍集合论与近世代数基础。在集合论部分，包括集合的概念、集合的运算与集合的重要性质；关于近世代数部分，重点放在群的结构上。特别强调如何由已知的群生成新的群（子群、积群、商群），以及这些新生群的运算结构。通过生成新群的思路，显示出近代应用数学的思维方法，这在本课程中是重中之重。

第 2 章是线性空间与线性变换。一方面是高等数学中线性代数的继续，另一方面是近世线性代数所涵盖的一部分数学基础，涉及正交几何与辛几何的结构，这在近代数学、近代物理与天文学中都起着重要作用。此外，从线性算子空间的高度，进一步认识线性空间与线性变换；从多重线性代数的角度认识张量空间等，都为学生理解与应用近代数学打下坚实的基础。

第 3 章集中介绍点集拓扑知识。集合的拓扑结构与集合的运算结构相平行，共同构成近代数学对集合的重要处理方式，也是近代数学的重要思维方法之一。具有拓扑结构的集合称为拓扑空间。拓扑的概念来自实践，来自欧氏空间，但它高度抽象，不仅蕴含了精致的数学思维，而且成为近代应用数学中必不可少的重要概念之一。这里也要特别强调两点：一是如何从已知的拓扑空间生成新的空间（子空间、积空间、商空间），这些新生空间的拓扑

结构，是读者应当注意的重点；二是空间上的连续映射与空间的紧性，请读者特别注意它们的意义与作用。

第 4 章是泛函分析基础。近代科学读物中常常见到的度量空间、Banach 空间、Hilbert 空间等，在本章中不仅给出确切的定义，而且给出它们的重要性质。由于近代数学有高度的抽象性，许多定理与性质的证明都具有清晰的思路与巧妙的方法，这不是我们在近代数学基础中所要强调的。但是对于一些具有代表性的证明，例如度量空间的完备化定理、开映射定理、共鸣定理等，我们仍不惜用较多的篇幅给出，是因为这些证明中包含了丰富的近代数学思想。泛函分析中的算子理论部分，特别是关于算子序列的收敛性，也作了重点介绍。线性泛函部分则详细介绍了共轭空间与共轭算子。

第 5 章是分布理论，包含了 Fourier 变换的完整内容及其近代的发展 —— 小波变换。从 $L^p(\mathbb{R}^n)(1\leqslant p\leqslant 2)$ 的 Fourier 变换、Schwartz 函数的 Fourier 变换，到 Schwartz 分布的 Fourier 变换，详细介绍了它们的定义、性质与关系。Fourier 变换在近代科学技术中的作用毋庸置疑。数学家在 20 世纪中叶为 Fourier 分析奠定了新的基础 —— 分布理论。从分布理论的高度重新认识 Fourier 变换，以近代数学的观点与思维对待科学研究的对象，对每一个科学研究者至关重要。希望本章能起到这个作用。

第 6 章是流形上的微积分，重点介绍了光滑流形、余切空间、切空间的引入，流形上的微分与积分的定义及运算，并以三维欧氏空间作为辅助模型，给出各种应用举例。在概念方面，本章具有高度的抽象性，仅从定义去理解，有相当大的难度。因此，本章是本书的难点部分。我们按照数学大师陈省身先生引进微分流形的思路编写了本章的内容，力图使读者从实际的三维空间出发去进行一般情形的抽象。期望读者能领会从特殊到一般、从具体到抽象、从有限到无限、从理论到应用的学习方法，从而体会近代应用数学的真谛。

第 7 章的补充知识，包含了近代应用数学中常用的方法，如变分方法；常用的定理，如 Banach 空间中的 Stone-Weierstrass 定理、隐映射定理、逆映射定理、不动点原理；还有常用的概念，如 Haar 积分的概念。

本书所选用的知识载体跨度较大，几乎包含了非数学类自然科学领域的研究工作所必备的近代应用数学知识。在数年、数次的教学实践中，我们还着力于教学手段的改革，尽力以形象思维启发抽象思维，以几何直观引导分析概念，并辅以启发性的思考题，形成一个既可独立使用的完整课程教材，又可供某些系科参考的辅助教材。

本书曾由南京大学数学系孙永忠教授、邱华副教授作为教材使用过，他们都提出了有价值的意见与建议。匡亚明基础学院卢德馨教授、许旺教授以及物理系肖明文教授对本书提出了宝贵意见，于本书的使用、修改起了重要作用。清华大学出版社的石磊副编审对本书编写和出版提出了很好的建议，陈明编辑对书稿进行了细致核校，在此一并致谢。

由于作者水平有限，对教学改革的理解还不够深刻，教材改革的思路、取材、编写等方面的不足与错误在所难免，更存在许多不尽如人意之处，敬请专家、同行和读者不吝赐教。

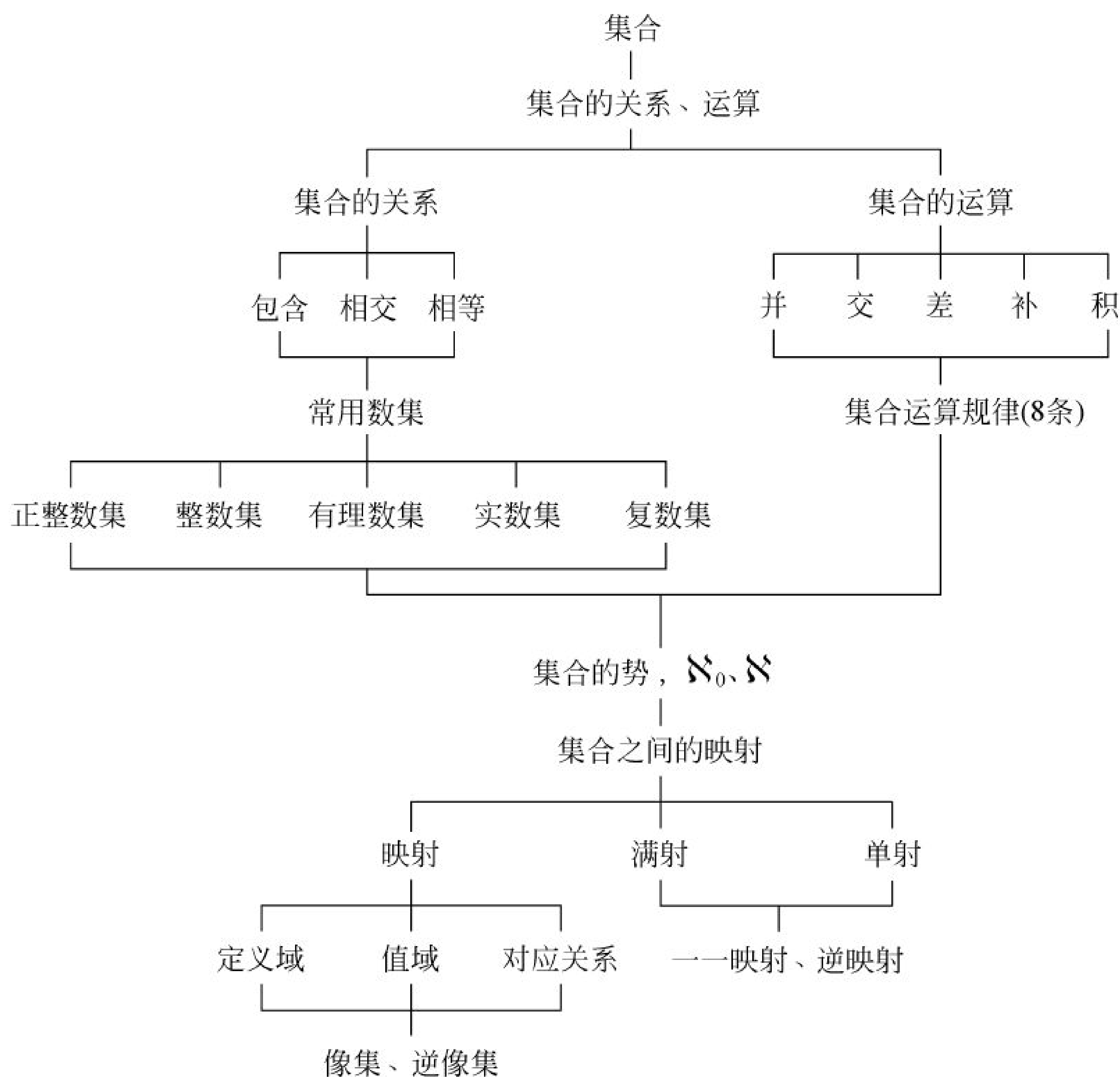
作　者

2015 年 1 月于南京大学

目　录

第1章 集合与集合的运算结构

集合论是近代数学的基础,我们从学习集合论的基本知识开始. 首先列出一张框图.



1.1 集合及其运算

集合论是 19 世纪后期由德国数学家 G. Cantor 创立的,经过一个多世纪,已经发展成为一个重要的数学分支;同时,也已经渗透到数学科学的各个领域,有大量的学术论文与专

著可供参考. 这里仅介绍集合论的基本知识. 本章主要参考文献为[1],[2].

1.1.1 集合

一个重要的数学思维方法是：把研究对象按照一定的性质分成各种类型进行研究，分在同一类型中的对象具有共同的性质，因此不需一个一个地考虑，而是按类集中，这样更能使我们了解一类事物的本质. 例如，把全体有理数归为一类，称为**有理数集**，记为

$$\mathbb{Q}=\left\{x: x=\frac{q}{p}, q\in\mathbb{Z}, p\in\mathbb{Z}^{+}\right\},$$

其中$\mathbb{Z}=\{\cdots,-2,-1,0,1,2,\cdots\}$为整数的全体，$\mathbb{Z}^{+}=\{1,2,\cdots\}$为正整数的全体.

又如，在区间$[a,b]$上连续函数的全体，称为$[a,b]$**上的连续函数集**，记为

$$C([a,b])=\{f:[a,b]\to\mathbb{R}, f(x)\text{ 是}[a,b]\text{ 上的连续函数}\};$$

$[a,b]$上具有一阶连续导数的连续函数的全体，称为$[a,b]$**上一阶连续可导函数集**，记为

$$C^{1}([a,b])=\{f(x):[a,b]\to\mathbb{R}, f(x), f'(x)\in C[a,b]\}.$$

一般地，把**具有一定性质的对象的全体称为一个集合(set)**(也简称**集**). 集合通常用大写英文字母 $A,B,\cdots$或 $X,Y,\cdots$表示. 集中的对象称为**元素(element)**(也简称**元**)，一般用小写英文字母 $a,b,\cdots$或 $x,y,\cdots$表示.

若 a 是 A 的元，便记为 $a\in A$；若 a 不是 A 的元，记为 $a\notin A$.

集合是一个非常有用的概念，更是很有用的数学工具，因为无论自然科学还是社会科学、人文科学，所研究的都是具有一定性质的对象，引进集合来描述这些对象，更便于用数学思维方法处理问题.

1. 常用的数集

正整数集(positive number set) $\mathbb{Z}^{+}=\{x: x=1,2,3,\cdots\}$,

自然数集(natural number set) $\mathbb{N}=\{x: x=0,1,2,3,\cdots\}$,

整数集(integer set) $\mathbb{Z}=\{x: x=0,\pm1,\pm2,\cdots\}$,

有理数集(rational number set) $\mathbb{Q}=\left\{x: x=\frac{q}{p}, q\in\mathbb{Z}, p\in\mathbb{Z}^{+}\right\}$,

实数集(real number set) $\mathbb{R}=\{x: x\text{ 为实数}\}$,

复数集(complex number set) $\mathbb{C}=\{z: z=a+\mathrm{i}b, a,b\in\mathbb{R}\}$, $\mathrm{i}=\sqrt{-1}$.

2. 集合之间的关系

包含(include) 设 A 与 B 为两个集合. 若 A 中任一个元 $a\in A$ 同时也属于 B，即 $x\in B$，则称 A **含于** B，记为 $A\subseteq B$，或称 B **包含** A，记为 $B\supseteq A$，并称 A **为** B **的子集(subset)**.

若 A 含于 B，但 $A\neq B$，则称 A 为 B 的**真子集(proper subset)**，记为 $A\subset B$.

例如，$\mathbb{Z}^+\subset\mathbb{N}\subset\mathbb{Z}\subset\mathbb{Q}\subset\mathbb{R}\subset\mathbb{C}$.

相等(equal) 若 $A\subseteq B$，且 $B\subseteq A$，则称 A **与** B **相等**，记为 $A=B$.

例如，$\{x\in\mathbb{N}:x^2=1\}=\{1\}\subset\{x\in\mathbb{Z}:x^2=1\}=\{1,-1\}$.

1.1.2 集合的运算

1. 集合的并、交、差、补、积

空集(empty set) 不含任何元素的集合称为空集，记为$\varnothing$.

例如$\{x\in\mathbb{N}:x^2=-1\}=\varnothing$，但是$\{z\in\mathbb{C}:z^2=-1\}\neq\varnothing$.

集合的并集(union set) 设 A 与 B 为两个集合. A 与 B 的并集(图 1.1.1(a))定义为一个集合，其中的元素或者属于 A，或者属于 B. 记为

$$A\cup B=\{x:x\in A\text{ 或 }x\in B\}.$$

集合的交集(intersection set) 设 A 与 B 为两个集合. A 与 B 的交集(图 1.1.1(b))定义为一个集合，其中的元素既属于 A，又属于 B. 记为

$$A\cap B=\{x:x\in A\text{ 且 }x\in B\}.$$

集合的差集(difference set) 设 A 与 B 为两个集合. A 与 B 的差集(图 1.1.1(c))定义为一个集合，其中的元素属于 A，但不属于 B. 记为

$$A\backslash B=\{x:x\in A\text{ 且 }x\notin B\}.$$

集合的补集(complementary set) 设 X 是基本集合，$A\subset X$ 为其子集. 称差集 $X\backslash A$ 为 A 关于基本集合 X 的补集(图 1.1.2(a)). 记为 $A^{C}=X\backslash A$，或 $\complement A$.

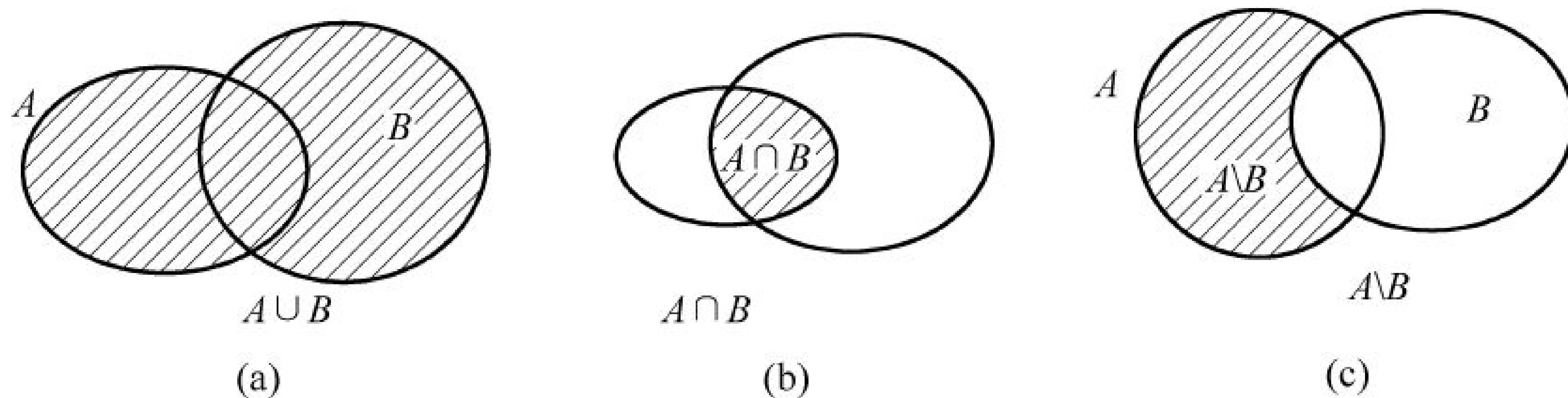


图 1.1.1 集合的并、交、差

集合的积集(product set) 设 A 与 B 为两个集合. A 与 B 的积集(图 1.1.2(b))定义为一个集合，记为

$$A\times B=\{(x,y):x\in A,y\in B\}.$$

n 个集合的积集记为

$\prod_{j=1}^{n}A_j\equiv A_1\times A_2\times\cdots\times A_n=\{(x_1,x_2,\cdots,x_n):x_j\in A_j,j=1,2,\cdots,n\}$. 也可以考虑

无限多个集合的积集$\prod_{\lambda\in\Lambda}A_\lambda$,这里指标集$\Lambda$可以是可数集,如$\prod_{j\in\mathbb{N}}A_j$,也可以是任意集.

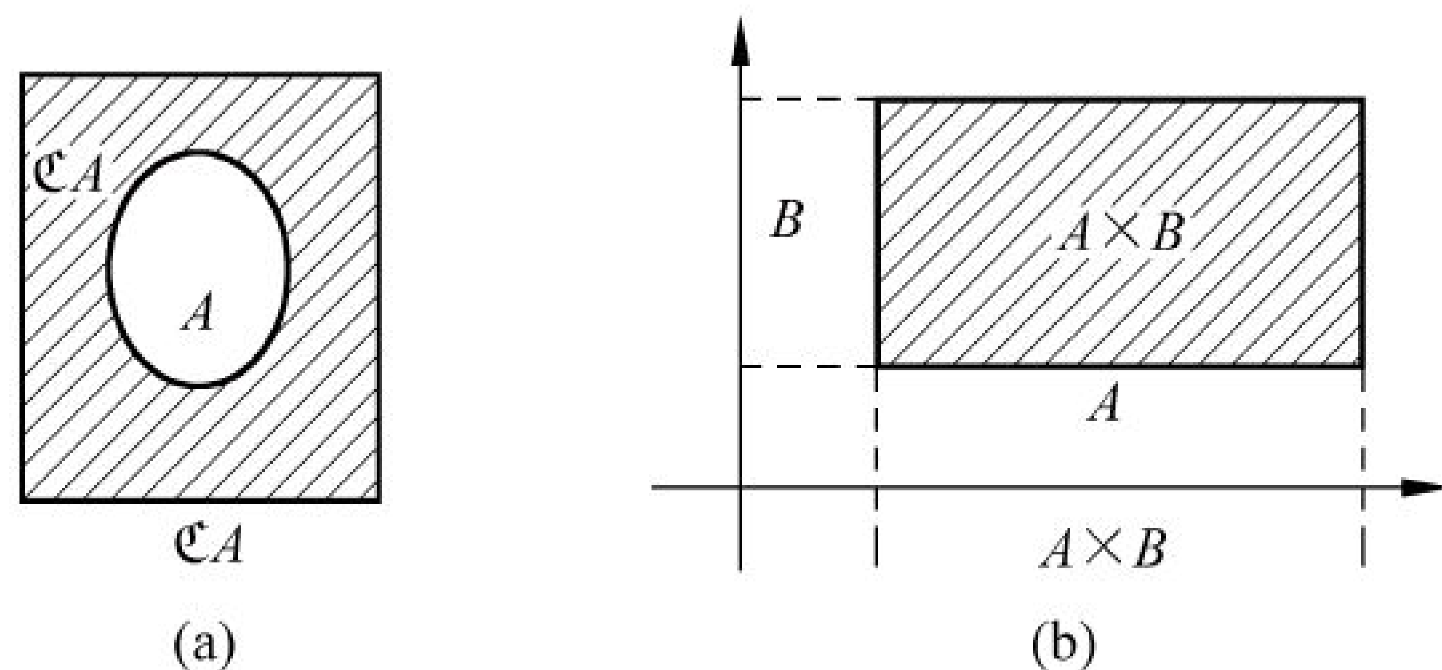


图1.1.2 集合的补、积

2. 集合的运算律

(1) $A\cap B\subset A\subset A\cup B$, $A\cap B\subset B\subset A\cup B$; (包含律)

(2) $A\cap B=B\cap A$, $A\cup B=B\cup A$; (交换律)

(3) $A\cap(B\cap C)=(A\cap B)\cap C$, $A\cup(B\cup C)=(A\cup B)\cup C$; (结合律)

(4) $A\cap(A\cup B)=A\cup(A\cap B)=A$; (吸收律)

(5) $A\cap(B\cup C)=(A\cap B)\cup(A\cap C)$, $A\cup(B\cap C)=(A\cup B)\cap(A\cup C)$; (分配律)

(6) $\complement(\complement A)=A$, $\complement(A\cap B)=\complement A\cup\complement B$, $\complement(A\cup B)=\complement A\cap\complement B$. (对合律)

还可以考虑有限多个集合或无限多个集合的并集与交集:

$$\bigcup_{j\in\Lambda}A_j,\quad \bigcap_{j\in\Lambda}A_j,$$

这里指标集Λ可以是有限集或无限集,并且有以下性质:

(7) $A\cap\left(\bigcup_{j\in\Lambda}A_j\right)=\bigcup_{j\in\Lambda}(A\cap A_j)$, $A\cup\left(\bigcap_{j\in\Lambda}A_j\right)=\bigcap_{j\in\Lambda}(A\cup A_j)$;

(8) $\complement\left(\bigcup_{j\in\Lambda}A_j\right)=\bigcap_{j\in\Lambda}\complement A_j$, $\complement\left(\bigcap_{j\in\Lambda}A_j\right)=\bigcup_{j\in\Lambda}\complement A_j$. (de Morgan 公式)

1.1.3 集合之间的映射

函数是高等数学中的重要概念之一,高等数学中函数f的定义域D_f是实数集$\mathbb{R}$中的集合(一元函数),或是$\mathbb{R}^n$中的集合(多元函数);值域R_f总是在$\mathbb{R}$中.然而,近代科学技术中所遇到的量的变化范围、量与量的关系远远超出了欧氏空间$\mathbb{R}$与$\mathbb{R}^n$.因此,在集合上引进“对应关系”,以刻画变量之间的关系,也就是必然的了.

1. 映射

定义1.1.1(对应关系) 设X,Y为两个集合(可以相同,也可不同).对于给定的两个子集$A\subseteq X$,$B\subseteq Y$,如果有一个**对应关系(corresponding relation)**,记为$f:A\to B$,使得对每

一个 $x\in A$,有**惟一的**(**unique**)$y\in B$ 与之对应,记为 $y=f(x)$,则称 $f:A\to B$ 为 A 到 B 的**映射**(**mapping**),或**变换**(**transform**),或**算子**(**operator**). 称 A 为 f 的**定义域**(**domain**),记为 $\mathfrak{D}$;称 B 为 f 的**值域**(**range**),记为 $\mathfrak{R}$. 称 $f(A)=\{f(x):x\in A\}\subseteq B$ 为 f 的**像集**(**image set**),记为 $\mathrm{im}(f)=f(A)$;对子集 $C\subseteq B$,称集 $f^{-1}(C)=\{x\in A:f(x)\in C\}\subseteq A$ 为 f 的**逆像集**(**inverse image set**).

易见,$f(A)\subseteq B$ 有可能是 $f(A)=B$,也有可能是 $f(A)\subset B,f(A)\neq B$.

满射(**surjective**) 当 $f(A)=B$ 时,称 f 为 A 到 B 的满射. 亦即,对每个 $y\in B$,至少存在一个 $x\in A$,满足 $f(x)=y\in Y$. 此时,f 的像集 $f(A)$ 充满整个 B,即 $\mathrm{im}(f)=B$.

单射(**injective**) 若映射 $f:A\to B$ 是 A 到 B 的一一对应映射,即"$f(x_1)=f(x_2)$ 蕴含 $x_1=x_2$",就称映射 f 为单射,或称**一一**(**one to one**)**映射**.

注意,一一映射 $f:A\to B$ 未必是满射.

逆映射(**inverse**) 对于一一映射(单射)$f:A\to B$,定义其**逆映射** $f^{-1}:B\to A$ 为满足"$f(x)=y$ 蕴含 $f^{-1}(y)=x$"的映射 f^{-1}. (请读者区分逆映射 $f^{-1}:B\to A$ 与逆像集 $f^{-1}(B)\subset X$.)

例 1.1.1 设 $A=\{a,b,c\},B=\{1,2,3,4\}$. 映射 $f:A\to B$ 满足

$$f(a)=1,\quad f(b)=2,\quad f(c)=3,$$

则 f 是 A 到 B 的一一映射,但不是满射.

例 1.1.2 取 $A=B=\mathbb{Z}=\{0,\pm1,\pm2,\pm3,\cdots\}$,映射 $f:A\to B$ 满足

$$f(n)=n+1,$$

则 f 是$\mathbb{Z}$到$\mathbb{Z}$的一一映射,并且是满射.

定义 1.1.2(**映射的复合**) 设 A,B,C 是三个给定的集合,映射 $f:A\to B,g:B\to C$,则由映射 f 与 g 确定的、$x\in A$ 到 $z\in C$ 的、满足

$$z=h(x)=g(f(x))$$

的映射 $h:A\to C$,称为**映射 f 与 g 的复合**(**compound**),记为 $h=g\circ f$.

注意,在考虑映射的复合时,像集 $f(A)$ 必须包含在映射 $g:B\to C$ 的定义域 B 中,但未必一定等于 B.

例 1.1.3 设 $A=C[a,b]$. 令 $J:C[a,b]\to\mathbb{R}$ 是由下面的积分确定的映射:

$$J(f)=\int_a^b f(x)\mathrm{d}x,$$

这里 $J(f)$ 为 f 的黎曼积分. 当 $f\in C[a,b]$时,$J(f)$视为 f 的函数,称 $J(f)$为集合 $C[a,b]$ 上的一个泛函. 下一节引入空间概念后,对于泛函还要给出更确切的定义.

2. 集合的势

"集合中元素个数"表示集合中元素的多少. 在自然科学问题的研究中,集合所含元素

的个数往往反映集合的许多方面的性质,因此非常重要. 但是,“多少”这个概念,对于“有限多个”而言才是有意义的,对于“无限多个”,就失去了意义.

有限集(finite set) 若集合 A 中含有限多个元素,即

$$A=\{a_1,a_2,\cdots,a_n\},\quad n\in\mathbb{Z}^+,$$

这里 $n\in\mathbb{Z}^+$ 是一个确定的、有限的正整数,则称 n 是 A 的**个数**,并称 A 为**有限集**.

无限集(infinite set) 若 A 的个数不是有限的,即 A 不是有限集,则称 A 为**无限集**,也称**无穷集**.

例如,$\{1,2,3,4\}$,$\mathbb{Z}_2=\{0,1\}$都是有限集,而$\mathbb{Z}^+$,$\mathbb{N}$,$\mathbb{Z}$,$\mathbb{Q}$,$\mathbb{R}$,$\mathbb{C}$都是无限集.

显然,对无限集谈“个数”是没有实际意义的. 例如,自然数集$\mathbb{N}=\{0,1,2,\cdots\}$与偶数集$\mathbb{O}=\{2,4,6,\cdots\}$都含有无限多个元素,都是无限集,看似自然数集中的元比偶数集的元多出一倍,但偶数也是非常非常多的,因此说哪个集合中的元素更多,没有实际意义.

为了比较两个无限集元素的“多少”,我们注意到两个相等的有限集 A 与 B 有一个重要的**特征性质**,那就是,两者的元素之间可以建立一一对应关系. 正是这种特征性质的启发,使得我们用下面的方法来研究无限集中含元素的多少.

定义 1.1.3(对等) 若集合 A 与 B 之间存在对应关系 $\varphi:A\to B$,使得 φ 是 A 到 B 的单射,同时也是满射,亦即,φ 是 A 到 B 上的一一映射,且 $B=\varphi(A)$,则称集合 A 与 B **对等(equivalence)**,记为 $A\sim B$,称 φ 为 A 到 B 上的**对等关系(equivalence relation)**.

根据定义,**两个有限集是对等的,当且仅当它们的个数相等.**

但是,无限集就完全不同了.

正整数集$\mathbb{Z}^+=\{1,2,\cdots\}$与偶数集$\mathbb{O}=\{2,4,6,\cdots\}$中元素相比较,正整数要比偶数多一倍,但是,它们却是对等的(请读者给出对等关系 φ). 又如,有理数集$\mathbb{Q}$与正整数集$\mathbb{Z}^+$,都含有无限多个元素,而有理数集中所含的元素,要比正整数集中的元素多得多,但$\mathbb{Q}$与$\mathbb{Z}^+$也是对等的(请读者给出对等关系 φ).

定义 1.1.4(无限集的势) 若两个无限集合 A 与 B 是对等的,则称它们具有相同的**势(cardinal number)**,集合 A 的势记为 $\overline{\overline{A}}$. 故对等集 A 与 B 满足 $\overline{\overline{A}}=\overline{\overline{B}}$.

有限集的个数、无限集的势,统称为集合的**基数(cardinality)**,记为 $\mathrm{card}(A)=\overline{\overline{A}}$.

无限集与它的子集的势,是可以比较大小的.

势的比较 对于无限集 B 的一个真子集 $A\subset B$,$A\neq B$,有

(1) 若 A 为有限集,则 $\overline{\overline{A}}<\overline{\overline{B}}$;

(2) 若 A 为无限集,且 $A\sim B$,则 A 的势等于 B 的势,即 $\overline{\overline{A}}=\overline{\overline{B}}$;

(3) 若 A 为无限集,且 A 与 B 不对等,则 A 的势小于 B 的势,$\overline{\overline{A}}<\overline{\overline{B}}$;或说 B 的势大于 A 的势,记为 $\overline{\overline{B}}>\overline{\overline{A}}$.

在数的集合中,正整数集$\mathbb{Z}^+$是最小的无限集,称它的势为$\aleph_0$,读作**阿列夫零**. 势为$\aleph_0$的集称为**可数集**. 凡与$\mathbb{Z}^+$对等的无限集的势,都是$\aleph_0$. 显然,$\mathbb{Z}^+\sim\mathbb{N}\sim\mathbb{Z}\sim\mathbb{Q}$,因此$\mathbb{N}$,$\mathbb{Z}$与$\mathbb{Q}$的势都是$\aleph_0$.

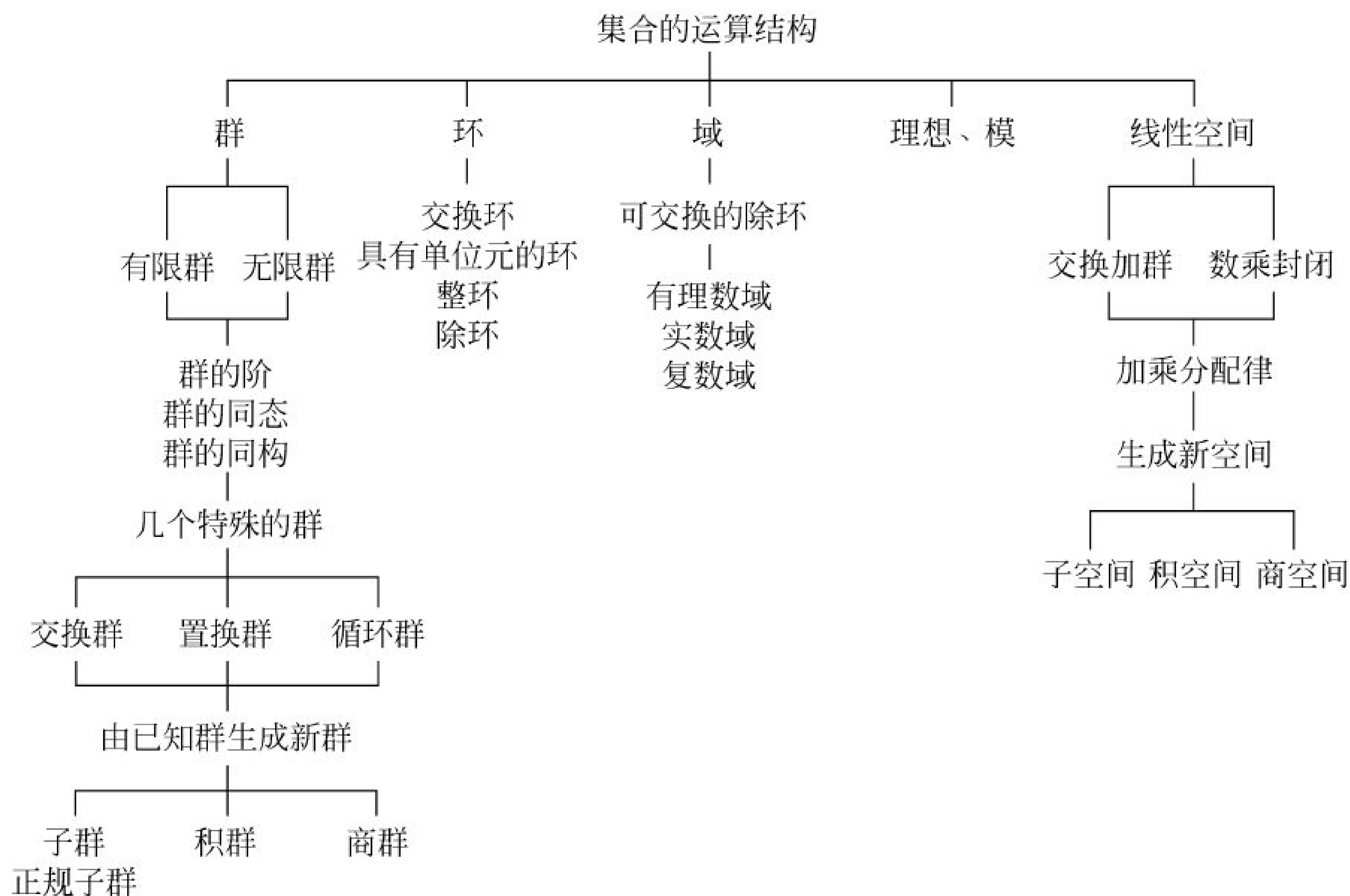
$[0,1]$中含有无限多个数,包括有理数与无理数,它不可能与$\mathbb{Z}^+$建立一一对应(这在实变函数教程中有详细证明,参考[6]、[14]),称$[0,1]$的势为$\aleph$,读作**阿列夫**. 于是,$\aleph_0<\aleph$.

具有势$\aleph$的数集有$[a,b]$,$[a,b)$,$(a,b]$,(a,b),$\mathbb{R}$等.

1.2 集合的运算结构

考虑集合中元素间的关系,有两个重要方面,一是元素之间的运算关系,如元素之间的加法、乘法等;二是元素间的位置关系,如两个元素间的距离,即远近的程度. 当然还有其他方面的关系,读者会在以后的学习或研究中遇到.

本节介绍集合中元素的一些运算关系,赋予代数运算的集合称其具有**运算结构**,也常称其具有代数结构. 本节内容见下图示意.



1.2.1 群、环、域、线性空间

1. 群

先从大家熟悉的实数集$\mathbb{R}$开始. 对于$x,y\in\mathbb{R}$,和式$x+y$可看作运算$+$的结果. 这种运算有以下熟知的性质:

封闭性 $x,y\in\mathbb{R}\Rightarrow x+y\in\mathbb{R}$；

结合律 $x,y,z\in\mathbb{R}\Rightarrow(x+y)+z=x+(y+z)$；

单位元 存在单位元 $0\in\mathbb{R}$，使得 $\forall x\in\mathbb{R}\Rightarrow x+0=0+x=x$；

逆元 $\forall x\in\mathbb{R}$，存在逆元 $(-x)\in\mathbb{R}\Rightarrow x+(-x)=0$.

这时，称$\mathbb{R}$在运算+之下构成一个群.

把这个定义抽象化，得到群的概念.

定义 1.2.1(群) 如果在一个集合 G 上定义一个运算，记为·，满足

(1) $x,y\in G\Rightarrow x\cdot y\in G$； (封闭性)

(2) $x,y,z\in G\Rightarrow(x\cdot y)\cdot z=x\cdot(y\cdot z)$； (结合律)

(3) 存在单位元 $1\in G$，使得 $\forall x\in G\Rightarrow 1\cdot x=x\cdot 1=x$； (单位元)

(4) $\forall x\in G$，存在逆元 $x^{-1}\in G\Rightarrow x\cdot x^{-1}=x^{-1}\cdot x=1$， (逆元)

则称 G 在运算·之下构成一个**群(group)**，1 称为群的**单位元(unit)**. 如果运算·还满足

(5) $x,y\in G\Rightarrow x\cdot y=y\cdot x$， (交换律)

则称 G 在运算·之下构成一个**交换群(exchange group)**，也称为 **Abel 群**.

群 G 与其运算·，常记为 $(G,\cdot)$，或简记为 G.

若集合 G 上的运算·仅满足定义 1.2.1 中的(1)、(2)，则称 $(G,\cdot)$ 为一个**半群(semi-group)**.

定义 1.2.1 中定义的运算·的意义是很广泛的，可取作实数的加法、乘法，也可取作函数的复合等. 实数集 $\mathbb{R}=\{x:-\infty<x<+\infty\}$ 在数的加法+之下构成一个交换群 $(\mathbb{R},+)$，其单位元就是数 0. 今后，在不发生混淆时，省去运算符号·.

例 1.2.1 记

$$\mathbb{R}^{+}=\{x\in\mathbb{R}:x>0\}$$

为**正实数集合**，把运算·取作数的乘法×. 不难验证，$\mathbb{R}^{+}$ 在运算×之下成为一个交换群 $(\mathbb{R}^{+},\times)$(请读者自行验证)，它的单位元是 1. 也请读者考虑，$(\mathbb{R},\times)$ 是否能构成一个群？$(\mathbb{R}\setminus\{0\},\times)$ 是否能构成一个群？

例 1.2.2 取

$$C([a,b])=\{f:[a,b]\to\mathbb{R},f(x)\text{ 是}[a,b]\text{ 上的连续函数}\}.$$

把运算·取作函数的加法+，即 $(f+g)(x)=f(x)+g(x)$，则 $(C([a,b]),+)$ 构成一个可交换的加群.

例 1.2.3 设

$$L(\mathbb{R},\mathbb{R})=\{T:T(x)=ax+b,a,b\in\mathbb{R},a\neq 0\}$$

为$\mathbb{R}$到$\mathbb{R}$的线性映射的全体，$T(x)$ 称为**仿射变换(affine transformation)**. 把运算·取作函数的复合∘，则

$$(L(\mathbb{R},\mathbb{R}),\circ)$$

构成一个群.

事实上，对于 $S,T\in L(\mathbb{R},\mathbb{R})$，令

$$S(x)=ax+b,\quad T(x)=cx+d,\quad a\neq 0,c\neq 0.$$

则

$$(T\circ S)(x)=T(S(x))=c(S(x))+d=c(ax+b)+d=(ca)x+(cb+d),$$

因 $a,c\neq 0$，故 $T\circ S\in L(\mathbb{R},\mathbb{R})$，此即定义 1.2.1 中的封闭性(1).

定义 1.2.1 中的结合律(2)显然成立.

对于定义 1.2.1 中的(3)，单位元为恒同映射 $I:x\rightarrow x,I(x)=x$，亦即 $a=1,b=0$.

对于定义 1.2.1 中的(4)，$\forall S(x)=ax+b\in L(\mathbb{R},\mathbb{R}),a\neq 0$，其逆元 S^{-1} 为

$$S^{-1}(x)=\frac{x}{a}-\frac{b}{a}\in L(\mathbb{R},\mathbb{R}).$$

故 $L(\mathbb{R},\mathbb{R})$是一个群，但一般不满足定义 1.2.1 中的(5)，故它不是交换群.

例 1.2.4 $(\mathbb{R},+),(\mathbb{Q},+),(\mathbb{Z},+)$都是加群，这里加法运算+就是实数的加法，它们都是加法交换群.

例 1.2.5 (1) $(\mathbb{R}\setminus\{0\},\times),(\mathbb{R}^+,\times)$都是乘法交换群，这里乘法运算×就是实数的乘法.

(2) 正整数乘法半群$(\mathbb{Z}^+,\times)$.

(3) 正整数的 p 倍半群$(\mathbb{Z}_p^+,\times)=(\{np:n\in\mathbb{Z}^+\},\times)$.

(4) p-adic 有限群$(\{0,1,\cdots,p-1\},\oplus)$，$p\geqslant 2$ 为正整数，这里$\oplus$是模 p 运算，$x\oplus y=x+y(\mathrm{mod}\,p)$.

(5) Cayley 有限群 $G=(\{1,\mathrm{i},-1,-\mathrm{i}\},\otimes)$，其中运算$\otimes$由下面的 Cayley 表给出：

$\otimes$	1	i	−1	−i
1	1	i	−1	−i
i	i	−1	−i	1
−1	−1	−i	1	i
−i	−i	1	i	−1

(6) 非异 n 阶复方阵群(复完全线性群)

$$(GL(n,\mathbb{C}),\times)=(\{A=[\alpha_{jk}]_{n\times n}:\alpha_{jk}\in\mathbb{C},\det A\neq 0\},\times),$$

其中×为方阵乘法，$\det A$ 是方阵 A 的行列式；

非异 n 阶实方阵群(实完全线性群)

$$(GL(n,\mathbb{R}),\times)=(\{A=[\alpha_{jk}]_{n\times n}:\alpha_{jk}\in\mathbb{R},\det A\neq 0\},\times);$$

当 $n\geqslant 2$ 时，它们都是非交换群. 且 $(GL(n,\mathbb{R}),\times)\subset(GL(n,\mathbb{C}),\times)$.

(7) 行列式为 1 的非异 n 阶复方阵群

$$(SL(n,\mathbb{C}),\times)=(\{A\in GL(n,\mathbb{C}):\det A=1\},\times);$$

行列式为1的非异 n 阶实方阵群

$$(SL(n,\mathbb{R}),\times)=(\{A\in GL(n,\mathbb{R}):\det A=1\},\times),$$

当 $n\geqslant 2$ 时,它们都是非交换群. 且 $(SL(n,\mathbb{R}),\times)\subset(SL(n,\mathbb{C}),\times)$.

(8) 正方形 $ABCD$ 绕其中心点将自身映射到自身的旋转群

$$\mathfrak{R}=\left\{\alpha:\alpha=0,\frac{1}{2}\pi,\pi,\frac{3}{2}\pi\right\},$$

其中 α 为旋转角,运算$\circ$定义为:$\alpha_1\circ\alpha_2$ 表示先旋转 α_2,再旋转 α_1,则$(\mathfrak{R}\circ)$构成旋转群.

(9) 单位圆根群

$$\Omega_p=\left\{\exp\left(\frac{\mathrm{i}2k\pi}{p}\right):k=0,1,\cdots,p-1\right\},$$

运算定义为复数乘法 $\times$,$p\geqslant 2$ 为正整数,则$(\Omega_p,\times)$构成单位圆根群.

例 1.2.6(子群) 给定群$(G,\cdot)$的子集 $H\subset G$,若在 G 的运算 $\cdot$ 之下,H 也成为一个群,则称 H 为 G 的子群.

例如,$GL(n,\mathbb{R})$是$GL(n,\mathbb{C})$的子群,$SL(n,\mathbb{R})$是$SL(n,\mathbb{C})$的子群.

2. 环、域

定义 1.2.2(环、域) 若在集合 R 上定义两个运算,分别记为加法$+$与乘法$\times$.

1) **环** —— 若 R 上的运算满足

(1) R 关于加法$+$构成一个交换群,其加法单位元记为0,称为 R 的**零元**;

(2) R 关于乘法$\times$构成一个半群,即满足封闭性与结合律,

$$x,y\in R\Rightarrow x\times y\in R;\qquad x,y,z\in R\Rightarrow (x\times y)\times z=x\times(y\times z);$$

(3) R 关于加法$+$与乘法$\times$满足加乘分配律,即

$$x,y,z\in R\Rightarrow x\times(y+z)=x\times y+x\times z,\quad (y+z)\times x=y\times x+z\times x;$$

则称 R 在运算$+$与$\times$之下构成一个**环(ring)**,记为$(R,+,\times)$.

2) **交换环** —— 若环$(R,+,\times)$还满足

(4) R 关于乘法$\times$满足交换律,即

$$x,y\in R\Rightarrow x\times y=y\times x,$$

则称$(R,+,\times)$为一个**交换环(exchange ring)**,亦即交换环满足(1)~(4).

3) **具有单位元的环** —— 若环$(R,+,\times)$还满足

(5) R 对乘法$\times$存在单位元,记为1,即

$$\forall x\in R\Rightarrow 1\times x=x\times 1=x,$$

则称$(R,+,\times)$为一个**具有单位元的环(ring with unit)**,亦即具有单位元的环满足(1)~(3)、(5).

注意,一个有单位元的环$(R,+,\times)$中,每个元未必有关于乘法的逆元,例如整数环$(\mathbb{Z},+,\times)$中,只有$+1$与-1关于乘法有逆元,其余整数关于乘法都没有逆元.

4) **整环** —— 若具有单位元的交换环$(R,+,\times)$还满足

(6) R没有零因子,

则称$(R,+,\times)$为**整环(integer ring)**,亦即整环满足(1)~(6).

(环R的**零因子(zero factor)**定义为:若对环$(R,+,\times)$中的元$a\neq 0$与$b\neq 0$,有$a\times b=0$,则称a为b的左零因子,称b为a的右零因子.若R为交换环,则左、右零因子统称为零因子.)

5) **除环(体)** —— 若具有单位元的环$(R,+,\times)$还满足

(7) R至少包含一个不等于零的元;

(8) R的每个非零元$x(x\neq 0)$都有逆元,

则称$(R,+,\times)$为**除环(division ring)**. 亦即除环满足(1)、(2)、(3)、(5)、(7)、(8).

6) **域** —— 一个交换除环称为**域(field)**,记为$(F,+,\times)$.

可以证明以下命题.

命题 1.2.1 (1) 一个除环(亦称为体)R没有零因子;

(2) 一个除环R使得$R\backslash\{0\}$关于乘法构成一个群.

命题 1.2.2 (1) 一个域F没有零因子;

(2) 一个域F使得$R\backslash\{0\}$关于乘法构成一个交换群.

域的等价定义 若集合F上有两个运算$+$与$\times$,使得$(F,+)$与$(F\backslash\{0\},\times)$成为两个交换群,并且加法$+$与乘法$\times$满足加、乘分配律,则$(F,+,\times)$称为一个域.

环与域的实例很多,以下再列举几个例子.

例 1.2.7 $(\mathbb{R},+,\times)$是最熟知的环,同时它也是一个域. 这里的$+$与$\times$就是通常$\mathbb{R}$上的加法与乘法.

事实上,集合$\mathbb{Z}$,$\mathbb{Q}$,$\mathbb{R}$,$\mathbb{C}$在通常加法$+$与乘法$\times$运算之下构成环,且都是交换环.

例 1.2.8 在计算机科学中有广泛应用的、仅含两个元素的域$\mathbb{Z}_2=\{0,1\}$,称为**二进伽罗瓦域(dyadic Galois field)**,其加法与乘法可用如下列表法定义:

$\oplus$	0	1
0	0	1
1	1	0

$\otimes$	0	1
0	0	0
1	0	1

3. 理想、主理想

定义 1.2.3(理想) 如果环$(R,+,\times)$的子集$J\subset R$满足

(1) J对加法$+$构成一个交换群$(J,+)$,加法单位元(零元)为$0\in J$;

(2) 对$a\in J,b\in R$,有$a\times b\in J,b\times a\in J$;

则称J为R的一个**理想(ideal)**.

条件(1)等价于$a,b\in J\Rightarrow a-b\in J$.

环$(R,+,\times)$中由一个元素$a\in R$生成的理想记为$\langle a\rangle=\{ra:r\in R\}$，称为由元$a\in R$生成的**主理想**(**principal ideal**).

设$(R,+,\times)$是一个整环，若它的所有理想都是主理想，则称$(R,+,\times)$为**主理想整环**(**principal ideal integer ring**).

例 1.2.9 整数环$(\mathbb{Z},+,\times)$是主理想整环，因为$\mathbb{Z}$中的任一理想$J\subset\mathbb{Z}$都是由J中的最小正整数生成的.

例 1.2.10 所有系数在域F上的单变量多项式的集合记为$F[x]$，它是一个具有单位元的交换环，并且是一个整环，也是一个主理想整环.

4. 环上的模、主理想整环上的模

定义 1.2.4(**环上的模**) 设$(R,+,\times)$是具有单位元的交换环，称集合M为环$(R,+,\times)$上的**模**(**modulo**)，若M上有两种运算：

"加法"$+$ $(u,v)\in M\times M\Rightarrow u+v\in M$;

"模乘"$\cdot$ $(r,u)\in R\times M\Rightarrow r\cdot u\in M$ ($r\cdot u$简记为ru)，

满足

(1) M对$+$构成一个交换群$(M,+)$，加法单位元(零元)为$0\in M$;

(2) M对模乘$\cdot$满足加法模乘分配律与乘法模乘结合律，即对任意的$u,v\in M$与$r,s\in R$，都有

$$r\cdot(u+v)=r\cdot u+r\cdot v,\quad (r+s)\cdot u=r\cdot u+s\cdot u$$

与

$$(r\times s)\cdot u=r\times(s\cdot u),\quad 1\cdot u=u,$$

其中1是环$(R,+,\times)$的乘法的单位元.

例 1.2.11 数域$\mathbb{F}$($\mathbb{R}$或$\mathbb{C}$)上的模M就是线性空间;

例 1.2.12 整数环$\mathbb{Z}$上的模M就是Abel群.

例 1.2.13 主理想整环$F[x]$上的模M，记为$F[x]$-模.

5. 线性空间

另外一种重要的运算结构是线性空间，它包含两种运算. 我们可以从读者熟悉的向量空间出发. 把解析几何中的三维欧氏空间$\mathbb{R}^3$用以下集合的记号来表示：

$$\mathbb{R}^3=\{x=(x_1,x_2,x_3):x_1,x_2,x_3\in\mathbb{R}\}.$$

在$\mathbb{R}^3$中，元素之间的加法和数乘运算是大家熟悉的，即对于$x,y\in\mathbb{R}^3$，有

$$x+y=(x_1,x_2,x_3)+(y_1,y_2,y_3)=(x_1+y_1,x_2+y_2,x_3+y_3),$$

$$\alpha\cdot x=\alpha(x_1,x_2,x_3)=(\alpha x_1,\alpha x_2,\alpha x_3),\quad \alpha\in\mathbb{R}.$$

不难验证，上述"加法"满足定义1.2.1，故$\mathbb{R}^3$成为一个可交换的加群；对于上述"数乘"还满足分配律与结合律. 由此推广到一般的集合上就得到下面定义.

定义 1.2.5(线性空间) 设 X 为一集合,其元素为 $x,y,\cdots$;$\mathbb{F}$ 为数域,可以是实数域 $\mathbb{F}=\mathbb{R}$ 或复数域 $\mathbb{F}=\mathbb{C}$. 如果对于 X 中任意两元素之间定义加法运算$+$与数乘运算 $\alpha\cdot$,使其具有如下性质:

(1) 加法运算$+$使得 X 成为一个**交换群**,亦即满足

① $x,y\in X \Rightarrow x+y\in X$; (封闭性)

② $x,y,z\in X \Rightarrow (x+y)+z=x+(y+z)$; (结合律)

③ X 中存在惟一的元 $0\in X$,使对任何 $x\in X$,有 $x+0=x$,称 0 为 X 的零元;

(存在加法单位元)

④ 对 X 中的任一元 $x\in X$,存在惟一的元$(-x)\in X$,满足 $x+(-x)=0$,并称$-x$ 为 x 的逆元; (存在加法逆元)

⑤ $x,y\in X \Rightarrow x+y=y+x$. (交换律)

(2) 数乘运算 $\alpha\cdot$ 满足

① $x\in X,\ \alpha\in\mathbb{F},1\in\mathbb{F}\Rightarrow \alpha\cdot x\in X,1\cdot x=x$; (封闭性)

② $x,y\in X,\ \alpha,\beta\in\mathbb{F}\Rightarrow \alpha\cdot(x+y)=\alpha\cdot x+\alpha\cdot y,\ (\alpha+\beta)\cdot x=\alpha\cdot x+\beta\cdot x$;

(分配律)

③ $x\in X,\alpha,\beta\in\mathbb{F}\Rightarrow (\alpha\beta)\cdot x=\alpha\cdot(\beta\cdot x)=\beta\cdot(\alpha\cdot x)$. (结合律)

则称 X 为**数域 $\mathbb{F}$ 上的线性空间**,简称**线性空间(linear space)**,或**向量空间(vector space)**,其元也称为**向量(vector)**.

以后在不引起混淆的情况下,我们常省略数乘运算中的$\cdot$.

显然$\mathbb{R}^3$ 为线性空间,亦称为三维欧氏空间(3-dimensional Euclidean space). 一般地,可把

$$\mathbb{R}^n=\{x=(x_1,x_2,\cdots,x_n):x_1,x_2,\cdots,x_n\in\mathbb{R}\}$$

定义为 n 维线性空间,或称 n 维欧氏空间(n-dimensional Euclidean space).

然而,具有线性空间结构的集合远远不止$\mathbb{R}^n$.

例 1.2.14 $C([a,b])=\{f:[a,b]\to\mathbb{R},f(x)$是$[a,b]$上的连续函数$\}$.

定义$[a,b]$上的两个连续函数 $f,g\in C[a,b]$的加法与数乘:

$$(f+g)(x)=f(x)+g(x),\quad x\in[a,b],$$

$$(\alpha f)(x)=\alpha f(x),\quad x\in[a,b],\ \alpha\in\mathbb{R},$$

则 $C[a,b]$成为一个线性空间,称为区间$[a,b]$上的**连续函数空间(continuous function space)**.

在本教程以后各章中会遇到更多重要的线性空间的例子.

1.2.2 群论初步、几种重要的群

群 $G\equiv(G,\cdot)$的定义已如定义 1.2.1 所述,视为一个代数系统,其运算$\cdot$称为乘法,意义是非常广泛的,只要满足定义 1.2.1 中所列条件,就使得集合 G 成为一个群,并称$(G,\cdot)$

具有**群结构**. 群结构只是运算结构的一种.

1. 有限群与无限群

群(G,·)中的元素个数可以是有限的,也可以是无限多的.

定义 1.2.6(有限群、无限群) 设 G 为一个群,若 G 中元素个数是一个有限的正整数,则称其为**有限群**,否则称为**无限群**.

1.2.1 节中的例 1.2.5 的(4)、(5)都是有限群. 它们的元素个数分别为 p 与 4.

无论是有限群还是无限群,都有以下重要的性质.

定理 1.2.1 设$(G,\cdot)$是一个群,则群的运算满足以下**消去律**:

(1) 对于 $a,x,x'\in G$,若 $ax=ax'$,则 $x=x'$;

(2) 对于 $a,y,y'\in G$,若 $ya=y'a$,则 $y=y'$.

证 只证(1). 由

$$ax=ax'\Rightarrow a^{-1}(ax)=a^{-1}(ax')\Rightarrow(a^{-1}a)x=(a^{-1}a)x'\Rightarrow x=x'$$

得证.

定义 1.2.7(群的阶,群的元 $a\in G$ 的阶) 设$(G,\cdot)$是一个群,若 G 是一个有限群,则其元的个数称为**群 G 的阶(degree of group G)**. 对于群的元 $a\in G$,设 e 为 G 的单位元,使得 $a^m=e$ 成立的最小的正整数 $m\in\mathbb{Z}^+$ 称为元 a 的**阶(degree of element a)**,并称 a 是 m 阶的. 若不存在这样的 m,则称 a 是**无限阶的**$\left(\text{当群运算为加法时},a^m\text{ 成为 }m\text{ 个 }a\text{ 相加,即 }\underbrace{a+a+\cdots+a}_{m}\right)$.

例 1.2.15 设 G 是由 $x^3=1$ 的三个根组成,运算×是数的乘法,则$(G,\times)$成为一个群.

解 由 $G=\left\{e=1,\varepsilon_1=\dfrac{-1+\sqrt{-3}}{2},\varepsilon_2=\dfrac{-1-\sqrt{-3}}{2}\right\}$,对于数的乘法,封闭性、结合律、单位元都是显然的,且 $\varepsilon_1^{-1}=\varepsilon_2,\varepsilon_2^{-1}=\varepsilon_1$. 于是,1 的阶为 $m=1$,ε_1 的阶为 $m=3$,ε_2 的阶为 $m=3$.

2. 群的同态、同构

定义 1.2.8(群的同态、同构) 设$(G,(\cdot)_G)$与$(H,(\cdot)_H)$是两个群,f 是 G 到 H 的映射. 如果对于 $\alpha,\beta\in G$,其像 $f(\alpha),f(\beta)\in H$,且满足

$$f(\alpha(\cdot)_G\beta)=f(\alpha)(\cdot)_H f(\beta),$$

则称 f 为群 G 到群 H 的**同态映射**,简称 f 为**同态(homomorphism)**. 如果 f 还是一一映上(onto)映射,即满射,则称 f 为 G 到 H 的**同构映射**,简称 f 为**同构(isomorphism)**.

显然,同态 $f:G\to H$ 就是保持群运算的映射.

例 1.2.16 大家熟悉的同构映射的例子是 $\log:\mathbb{R}^+\to\mathbb{R}$,它是群$(\mathbb{R}^+,\times)$到群$(\mathbb{R},+)$的同构,$\log(x\times y)=\log x+\log y$.

例 1.2.17 设 $G,\widetilde{G}$为两个群,$\cdot$ 与 $\tilde{\cdot}$ 分别为其运算,e 与 $\tilde{e}$分别为其单位元,令

$$f:G\to\widetilde{G},\quad f(x)=\tilde{e},\quad \forall x\in G,$$

则 f 是 G 到 $\widetilde{G}$的同态.

因为令 $z=x\cdot y\in G$,则有 $f(x\cdot y)=f(z)=\tilde{e}=\tilde{e}\ \tilde{\cdot}\ \tilde{e}=f(x)\ \tilde{\cdot}\ f(y)$.

例 1.2.18 设$(\mathbb{R},+)$与$(G,\times)$为两个群,这里 $G=\{z\in\mathbb{C}:|z|=1\}$为复平面$\mathbb{C}$上的单位圆周,它在运算$\times$之下成为一个群. 设映射 $\varphi:\mathbb{R}\to\mathbb{C}$ 为 $\varphi(x)=\mathrm{e}^{\mathrm{i}x}$,则 φ 是$\mathbb{R}$ 到$\mathbb{C}$ 上的同态.

定理 1.2.2 设 G 与$\widetilde{G}$是两个群,映射 $f:G\to\widetilde{G}$是一个满同态(满射、同态),则 G 的单位元 e 的像 $f(e)$是$\widetilde{G}$的单位元$\tilde{e}=f(e)$;任一个元 $a\in G$ 的逆元 a^{-1}的像 $f(a^{-1})$是 a 的像 $f(a)$的逆元 $f(a^{-1})=(f(a))^{-1}$.

证明留作习题.

究竟有多少种类型的群呢? 这是一个很复杂的问题. 本书的目的是介绍应用较为广泛的群,例如变换群(非交换的)、置换群(非交换的)、循环群.

3. 变换群

本小节给出一种常用的、元素是映射的群,它是不可交换群.

定义 1.2.9(集合的变换、变换群) 设 A 为一个集合,称由 A 到自身的映射 $\tau:A\to A$ 为集合 A 的一个变换. 记为 $a\in A\to\tau(a)\in A$.

设$\mathbb{S}=\{\tau:A\to A\}$为 A 到 A 自身的"**一一变换**"(既是单射又是满射)所成的集合,赋予$\mathbb{S}$一个运算$\circ$: $\forall\tau_1,\tau_2\in\mathbb{S}$,两个变换之间的运算定义为变换的复合$\circ$,亦即

$$\tau_1\circ\tau_2(a)=\tau_1(\tau_2(a)),\quad \forall a\in G.$$

于是,$(\mathbb{S},\circ)$构成一个群,称此群为集合 A 上的**变换群(transformation group)**.

集合 A 上的变换群一定存在. 例如,A 到自身的一一变换的全体所成的集合,对于变换的复合运算,满足封闭性、结合律、存在单位元、存在逆元,故成为 A 的一个变换群.

例 1.2.19 设 $A=\mathbb{R}^2$,平面上的两个坐标轴绕坐标原点 $O(0,0)$旋转一个角度 θ,记为 τ_θ,则$\mathbb{S}=\{\tau_\theta:\mathbb{R}^2\to\mathbb{R}^2\}$成为一个变换群.

证 对于 $\tau_\varphi,\tau_\psi,\tau_\chi\in\mathbb{S}$,有

(1) $\tau_\varphi\circ\tau_\psi=\tau_{\varphi+\psi}$;　　(2) $(\tau_\varphi\circ\tau_\psi)\circ\tau_\chi=\tau_\varphi\circ(\tau_\psi\circ\tau_\chi)$;

(3) $\tau_0\in\mathbb{S}$;　　(4) $(\tau_\theta)^{-1}=\tau_{-\theta}$.

例 1.2.20 设 $A=\mathbb{R}^2$,平面上的所有旋转 τ 与平移 t 所成的集合,成为 $A=\mathbb{R}^2$ 上的变

换群. 但此变换群却不是可交换的.

例如,取 $A=\mathbb{R}^2$ 的一个平移 t: 把(0,0)平移到(1,0),取旋转 τ: 绕原点旋转$\frac{\pi}{2}$. 则

$$t\circ\tau:(0,0)\to(1,0),\quad \tau\circ t:(0,0)\to(0,1),$$

于是,$t\circ\tau\neq\tau\circ t$.

变换群的重要性在于以下定理.

定理 1.2.3 任何群$(G,\cdot)$都同构于一个变换群.

证 设$(G,\cdot)$是一个群,$G=\{a,b,c,\cdots\}$,任取 $x\in G$,定义映射

$$\tau_x:g\in G\to xg\equiv\tau_x(g)\in G,$$

则它是 G 上的一个变换. 于是,$\forall x\in G$ 确定 G 上的一个变换. 记 $\widetilde{G}=\{\tau_x:x\in G\}$. 易见,映射

$$\varphi:x\in G\to\tau_x\in\widetilde{G}$$

是 G 到$\widetilde{G}$的满射. 进而,当 $x,y\in G$ 且 $x\neq y$ 时,由消去律 $xg\neq yg$,$\forall g\in G$,因此,映射 $\varphi:x\to\tau_x$ 也是单射,故它是一一映射.

再验证映射 $\varphi:x\to\tau_x$ 的同构性,即 $\tau_{x\cdot y}=\tau_x\circ\tau_y$. 任取 $g\in G$,推演如下:

$$\tau_{x\cdot y}(g)=(x\cdot y)g=x(yg)=x(\tau_y(g))=\tau_x(\tau_y(g))=(\tau_x\circ\tau_y)(g).$$

单位元的对应为 $\varphi:e\to\tau_e$. 于是得到 $G=\{a,b,c,\cdots\}\xleftrightarrow{\text{同构}}\widetilde{G}=\{\tau_x:x\in G\}$.

4. 置换群

置换群是变换群的一种,有很重要的理论与实际应用.

定义 1.2.10(置换、置换群) 设 A 为有限集,称 A 到自身上的一一映射 $\tau:A\to A$ 为一个**置换(permutation)**.

设 $A=\{a_1,a_2,\cdots,a_n\}$为含 $n\in\mathbb{Z}^+$ 个元素的有限集,A 的全体置换所成的集合,以两个置换 τ 与 ν 的复合 $\nu\circ\tau$ 作为乘法运算,则构成一个置换群,称为 n **置换群(n-permutation group)**,或称 n **次对称群(symmetric group of degree n)**,记为 S_n,即

$$S_n=\left\{\tau=\begin{pmatrix}1&2&\cdots&n\\k_1&k_2&\cdots&k_n\end{pmatrix}:1\leqslant k_j\leqslant n,k_i\neq k_j,\text{若 } i\neq j\right\},$$

这里 $\tau=\begin{pmatrix}1&2&\cdots&n\\k_1&k_2&\cdots&k_n\end{pmatrix}$表示 $\tau:1\xrightarrow{\text{换为}}k_1,2\xrightarrow{\text{换为}}k_2,\cdots,n\xrightarrow{\text{换为}}k_n$.

置换 τ 可写成

$$\tau=\begin{pmatrix}1&2&\cdots&n\\k_1&k_2&\cdots&k_n\end{pmatrix}=\begin{pmatrix}2&1&\cdots&n\\k_2&k_1&\cdots&k_n\end{pmatrix}=\cdots=\begin{pmatrix}n&2&\cdots&1\\k_n&k_2&\cdots&k_1\end{pmatrix},$$

都表示一个意思：将 1 换为 k_1，2 换为 k_2，……，n 换为 k_n，只是写法不同而已．例如，对于 $n=3$，S_3 中共 6 个元

$$\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix},\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix},\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix},\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix},\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix},\begin{pmatrix}1&2&3\\3&2&1\end{pmatrix},$$

而每个元都有 6 个表示，例如对于 $\tau=\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}$，有

$$\tau=\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}=\begin{pmatrix}1&3&2\\2&1&3\end{pmatrix}=\begin{pmatrix}2&1&3\\3&2&1\end{pmatrix}=\begin{pmatrix}2&3&1\\3&1&2\end{pmatrix}=\begin{pmatrix}3&1&2\\1&2&3\end{pmatrix}=\begin{pmatrix}3&2&1\\1&3&2\end{pmatrix}.$$

今后用 $\tau=\begin{pmatrix}1&2&\cdots&n\\k_1&k_2&\cdots&k_n\end{pmatrix}$$(1\leqslant k_j\leqslant n, k_i\neq k_j$，若 $i\neq j)$ 表示置换群 S_n 中的元．置换群的运算是两个置换的复合．$\tau_2\circ\tau_1$ 表示先作置换 τ_1，再作置换 τ_2．不难验证，S_n 在上述运算之下构成置换群．

置换群的运算 $\tau_2\circ\tau_1$ 规则　先写下置换 τ_1，将置换 τ_2 的第一行与置换 τ_1 的第二行写为一致，得到置换 τ_2 的新表示，然后得到复合的结果，即

$$\begin{aligned}\tau_2\circ\tau_1&=\begin{pmatrix}1&2&\cdots&n\\k_1&k_2&\cdots&k_n\end{pmatrix}\begin{pmatrix}1&2&\cdots&n\\j_1&j_2&\cdots&j_n\end{pmatrix}\\&=\begin{pmatrix}j_1&j_2&\cdots&j_n\\l_1&l_2&\cdots&l_n\end{pmatrix}\begin{pmatrix}1&2&\cdots&n\\j_1&j_2&\cdots&j_n\end{pmatrix}=\begin{pmatrix}1&2&\cdots&n\\l_1&l_2&\cdots&l_n\end{pmatrix}.\end{aligned}$$

例如，对于 S_3 的情形，$\tau_2\circ\tau_1=\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}$，因 $\tau_2=\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}=\begin{pmatrix}2&1&3\\3&1&2\end{pmatrix}$，于是

$$\tau_2\circ\tau_1=\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}=\begin{pmatrix}2&1&3\\3&1&2\end{pmatrix}\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}=\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix}.$$

注意，置换群是不可交换群．例如

$$\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}=\begin{pmatrix}2&1&3\\3&1&2\end{pmatrix}\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}=\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix},$$

$$\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}=\begin{pmatrix}1&3&2\\2&3&1\end{pmatrix}\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}=\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}.$$

定理 1.2.4　n 置换群 S_n 含有 $n!$ 个置换．

证　这个定理由初等数学中的排列即可得到．

命题 1.2.1　设 S_n 为 n 置换群，对于其中两种特殊置换

$$\tau_1=\begin{pmatrix}j_1&\cdots&j_k&j_{k+1}&\cdots&j_n\\j_1^{(1)}&\cdots&j_k^{(1)}&j_{k+1}&\cdots&j_n\end{pmatrix},\quad\tau_2=\begin{pmatrix}j_1&\cdots&j_k&j_{k+1}&\cdots&j_n\\j_1&\cdots&j_k&j_{k+1}^{(2)}&\cdots&j_n^{(2)}\end{pmatrix}$$

的乘积，有

$$\tau_2\circ\tau_1=\begin{pmatrix} j_1 & \cdots & j_k & j_{k+1} & \cdots & j_n \\ j_1^{(1)} & \cdots & j_k^{(1)} & j_{k+1}^{(2)} & \cdots & j_n^{(2)} \end{pmatrix}.$$

证 记

$$\tau_1:(j_1 \quad \cdots \quad j_k \quad j_{k+1} \quad \cdots \quad j_n)\to(j_1^{(1)} \quad \cdots \quad j_k^{(1)} \quad j_{k+1} \quad \cdots \quad j_n),$$

$j_{k+1},\cdots,j_n$ 不变，而且 $j_{k+1},\cdots,j_n$ 不能再是 $j_1,\cdots,j_k$ 的像，亦即，当 $i\leqslant k$ 时，只能有

$$j_i^{(1)}=j_l,\quad l\leqslant k.$$

因此，当 $i\leqslant k$ 时，有

$$\tau_2(\tau_1(j_i))=\tau_2(j_i^{(1)})=\tau_2(j_i)=j_i^{(1)};$$

当 $i>k$ 时，有

$$\tau_2(\tau_1(j_i))=\tau_2(j_i)=j_i^{(2)}.$$

故 $\tau_2\circ\tau_1=\begin{pmatrix} j_1 & \cdots & j_k & j_{k+1} & \cdots & j_n \\ j_1^{(1)} & \cdots & j_k^{(1)} & j_{k+1}^{(2)} & \cdots & j_n^{(2)} \end{pmatrix}$.

下面给出置换群的另一种表示.

定义 1.2.11（k 循环置换） 称置换群 S_n 中的置换 $\tau\in S_n$ 为 k **循环置换**（k-**recurring permutation**），若 τ 将 $(a_{i_1} \quad \cdots \quad a_{i_k} \quad a_{i_{k+1}} \quad \cdots \quad a_{i_n})$ 中的 a_{i_1} 变到 a_{i_2}，将 a_{i_2} 变到 a_{i_3}，……，将 $a_{i_{k-1}}$ 变到 a_{i_k}，将 a_{i_k} 变到 a_{i_1}，其余的不变. 将一个 k 循环置换记为 $(i_1 \quad i_2 \quad \cdots \quad i_k)$，或记为 $(i_1 \quad i_2 \quad \cdots \quad i_k)=(i_2 \quad i_3 \quad \cdots \quad i_k \quad i_1)=\cdots=(i_k \quad i_1 \quad \cdots \quad i_{k-1})$ 中的任一个.

任意的置换不一定是一个循环置换，例如，S_4 中的一个置换 $\tau=\begin{pmatrix}1&2&3&4\\2&1&4&3\end{pmatrix}$ 不是循环置换，因为 τ 使得每一个元都发生变动，并且 $1\to2,2\to1,3\to4,4\to3$，如果它是一个循环置换，则它一定是一个 4 循环置换的乘积. 但 $2\to1,4\to3$ 不是顺序的. 然而，$\tau=\begin{pmatrix}1&2&3&4\\2&1&4&3\end{pmatrix}$ 却是两个循环置换的乘积，因为

$$\tau=\begin{pmatrix}1&2&3&4\\2&1&4&3\end{pmatrix}=\begin{pmatrix}1&2&3&4\\1&2&4&3\end{pmatrix}\begin{pmatrix}1&2&3&4\\2&1&3&4\end{pmatrix}=(3\ 4)(1\ 2).$$

定义 1.2.12（不相连循环、对换） 称互相没有共同数字的循环置换为**不相连循环**；称 2 循环置换为**对换**.

一般地，我们有下面定理.

定理 1.2.5 n 置换群 S_n 中任一个置换 τ 都可写成若干个不相连循环的乘积；进而，每个循环置换可表示为若干个对换之乘积. 因此，每个置换都可表示为若干个对换之乘积.

证 用归纳法可得到，略.

例如 S_4 的全体共 $4!=24$ 个元，用循环置换写出如下：

(1);

(1 2),(3 4),(1 3),(2 4),(1 4),(2 3);

(1 2 3),(1 3 2),(1 3 4),(1 4 3),(1 2 4),(1 4 2),(2 3 4),(2 4 3);

(1 2 3 4),(1 2 4 3),(1 3 2 4),(1 3 4 2),(1 4 2 3),(1 4 3 2);

(1 2)(3 4),(1 3)(2 4),(1 4)(2 3).

又如 S_5:$\begin{pmatrix}1&2&3&4&5\\1&2&3&4&5\end{pmatrix}=(1)=(2)=(3)=(4)=(5)$;

$\begin{pmatrix}1&2&3&4&5\\5&1&2&3&4\end{pmatrix}=(1\ 2\ 3\ 4\ 5)=(2\ 3\ 4\ 5\ 1)=(3\ 4\ 5\ 1\ 2)=(4\ 5\ 1\ 2\ 3)=(5\ 1\ 2\ 3\ 4)$;

$\begin{pmatrix}1&2&3&4&5\\2&3&1&4&5\end{pmatrix}=(1\ 2\ 3)=(2\ 3\ 1)$.

…

定理 1.2.6 设 S_n 为 n 置换群,因每个置换都可表示为对换之乘积,故所用的对换个数的奇、偶性不变;亦即,一个置换或能分解成奇数个对换之乘积,或能分解为偶数个对换之乘积,二者必居其一.

定义 1.2.13(奇置换、偶置换) 设 S_n 为 n 置换群,若一个置换可表示为奇数个对换之乘积,则称其为**奇(odd)置换**;若一个置换可表示为偶数个对换之乘积,则称其为**偶(even)置换**.

奇置换与偶置换又可用下面方法表示.

将 n 个元的集合 $S=\{1,2,\cdots,n\}$ 上的置换群记为

$$\mathfrak{T}n) \equiv S_n = \{\sigma:\sigma\{1,2,\cdots,n\} = \{\sigma(1),\sigma(2),\cdots,\sigma(n)\}\},$$

显然,$\mathfrak{T}n)$ 中有 $n!$ 个元.

定义 $\mathfrak{T}n)$ 中的元 σ,τ 的运算:$\tau\circ\sigma$ 为两个置换的复合,即

$$(\tau\circ\sigma)(s) = \tau(\sigma(s)), \quad \forall s \in \mathfrak{T}n).$$

于是,在复合运算 $\tau\circ\sigma$ 之下,$(\mathfrak{T}n),\circ)$ 成为**置换群**,其单位元为恒同映射 $I(s)=s(s\in\mathfrak{T}n))$,并且有 $\sigma\in\mathfrak{T}(n)\Rightarrow\sigma^{-1}\in\mathfrak{T}n)$.

进而,$\forall\sigma\in\mathfrak{T}n)$,若 $\sigma=\begin{pmatrix}1&2&\cdots&n\\\sigma(1)&\sigma(2)&\cdots&\sigma(n)\end{pmatrix}$,作 $\sigma(j)-\sigma(k)$,让 $k<j$ 并让 k 与 j 跑遍 S,则

$$\prod_{1\leqslant k<j\leqslant n}(\sigma(j)-\sigma(k)) = \pm(2!)(3!)\cdots((n-1)!),$$

根据定义,若此积为正数,则 σ 为**偶置换**;若此积为负数,则 σ 为**奇置换**.

5. 循环群

定义 1.2.14(循环群) 若群 G 的每个元素都是 G 的某个固定元素 $a\in G$ 的乘方,亦即

$$\forall x \in G,\ \exists m \in \mathbb{Z}^+,\ \text{s.t. } x = a^m,$$

则称 G 为一个**循环群**(**circulate group**). a 称为 G 的**生成元**(**generator**),并将循环群记为 $G=(a)$.

例如,以模 p 加法运算所构成的 p-adic 群 $G=\{0,1,\cdots,p-1\}$ 就是一个循环群.

定理 1.2.7 循环群 $G=(a)$ 的结构完全由其生成元 a 的阶(定义 1.2.6)确定.

(1) 若 a 的阶为无限,则 G 与整数加群($\mathbb{Z}$,+)同构;

(2) 若 a 的阶为有限数 $m\in\mathbb{Z}^+$,则 G 与模 m 的加群($\mathbb{Z}_m$,$\oplus$)同构.

证 (1) 对于 a 的阶为无限的情形,$a^k=a^h \Leftrightarrow h=k$. 这是因为,当 $h=k$ 时,显然有 $a^k=a^h$. 反之,若 $a^k=a^h$,且 $h\neq k$,例如 $h>k$,于是 $a^k=a^h \Rightarrow e=a^{h-k}$,这与 a 的阶为无限相矛盾,故 $h=k$.

这样,φ:$a^k\to k$ 是 G 与加群 $\mathbb{Z}$ 之间的一一映射,并且 $a^h a^k=a^{h+k}\to h+k$,故 $G\leftrightarrow\mathbb{Z}$.

(2) 若 a 的阶为有限数 $m\in\mathbb{Z}^+$,即 $a^m=e$,则 $a^k=a^h \Leftrightarrow m\,|\,h-k$,这是因为,当 $m\,|\,h-k$ 时,$\exists q\in\mathbb{Z}^+$,使得 $h-k=mq \Rightarrow h=k+mq$,因此

$$a^h = a^{k+mq} = a^k a^{mq} = a^k\,(a^m)^q = a^k e^q = a^k;$$

反之,$a^h=a^k \Rightarrow h-k=mq+r$ 时,其中 $0\leqslant r\leqslant m-1$,则

$$e = a^{h-k} = a^{mq+r} = a^{mq}a^r = ea^r = a^r,$$

由阶的定义,得到 $r=0$,从而 $m\,|\,h-k$. 于是,$a^k=a^h \Leftrightarrow m\,|\,h-k$ 得证.

这样,$a^k\to[k]$,这里 $[k]\in\mathbb{Z}_m=\{0,1,\cdots,m-1\}$,$\mathbb{Z}_m$ 是一个模 m 加群,加法为模 m 加

$$\oplus:\ x\oplus y = x+y(\bmod m),$$

并且 $a^h a^k=a^{h+k}\to[h+k]=[h]+[k]$,从而 $G\leftrightarrow\mathbb{Z}_m=\{0,1,\cdots,m-1\}$.

定理 1.2.8 循环群 $G=(a)$ 是一个交换群.

请读者自行证明.

1.2.3 子群、积群、商群

在 1.2.2 节中给出的是一些具体的群,如变换群、置换群、循环群等. 本节要从已知的一个群出发,来构造新的、有理论意义及实际应用价值的群.

1. 子群、正规子群

定义 1.2.15(**子群、嵌入映射**) 设(G,·)是一个群,$S\subset G$ 是 G 的子集,若 S 在运算·之下也构成一个群,则称(S,·)为(G,·)的**子群**(**subgroup**).

在子群 $S\subset G$ 与群 G 之间,有一个重要的映射,称为**嵌入映射**(**embedding mapping**)I:$S\to G$,使得 $I(x)=x$,$\forall x\in S$. 亦即,$x\in S$ 映到它自身 $x\in G$ 的**恒同映射**(**identity mapping**).

定理 1.2.9 嵌入映射 $I:S\to G$ 是子群 $(S,\cdot)$ 到群 $(G,\cdot)$ 中(into)的同态.

证明留作习题.

例 1.2.21 $(\mathbb{R},+)$ 是一个群,$(\mathbb{Q},+)$,$(\mathbb{Z},+)$ 都是它的子群.

例 1.2.22 设 $G=S_3$,则 $H=\{(1),(1\quad 2)\}$ 是它的一个子群.

证 H 对 G 的乘法是封闭的.因为由

$$G=S_3=\{\sigma:\sigma(1),\sigma(2),\sigma(3)\},$$

即

$$S_3=\left\{\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix},\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix},\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix},\begin{pmatrix}1&2&3\\3&2&1\end{pmatrix},\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix},\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}\right\},$$

$$\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix}=(1)=(2)=(3),$$

$$\begin{pmatrix}1&2&3\\1&3&2\end{pmatrix}=(2\quad 3)=(3\quad 2),\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}=(1\quad 2)=(2\quad 1),\begin{pmatrix}1&2&3\\3&2&1\end{pmatrix}=(1\quad 3)=(3\quad 1),$$

$$\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix}=(1\quad 3\quad 2)=(2\quad 1\quad 3),\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}=(1\quad 2\quad 3)=(2\quad 3\quad 1).$$

注意,事实上也有

$$(1\quad 2)(1\quad 3)=(1\quad 2)\begin{pmatrix}1&2&3\\3&2&1\end{pmatrix}=\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix}=(1\quad 3\quad 2),$$

$$(1\quad 3)(1\quad 2)=(1\quad 3)\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}=\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}=(1\quad 2\quad 3).$$

子集 $H=\{(1),(1\quad 2)\}=\left\{\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix},\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}\right\}$ 有单位元 $(1)=\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix}$; 元 $(1\quad 2)=\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}$ 的逆元是其自身,因为

$$(1\quad 2)(1\quad 2)=(1\quad 2)\begin{pmatrix}1&2&3\\2&1&3\end{pmatrix}=\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix}=(1),$$

故 $H=\{(1),(1\quad 2)\}$ 是 S_4 的子群.

定理 1.2.10 群 G 的子集 $H\subset G$ 是子群,当且仅当

(1) $a,b\in H\Rightarrow ab\in H$;

(2) $a\in H\Rightarrow a^{-1}\in H$.

也可表示为:群 G 的子集 $H\subset G$ 是子群,当且仅当 $a,b\in H\Rightarrow ab^{-1}\in H$.

子群 H 的单位元就是群 G 的单位元;子群 H 的元 $a\in H$ 的逆元 $a^{-1}\in H$ 是 $a\in G$ 在 G 中的逆元.

定义 1.2.16(正规子群、陪集、中心) 设$(G,\cdot)$是一个群,若群G的子群$N\subset G$满足$\forall g\in G$,都有$gN=Ng$,则称N为G的**正规子群(normal subgroup)**,也称为**不变子群(invariant subgroup)**,并分别称$gN=\{gn:n\in N\}$与$Ng=\{ng:n\in N\}$为N的左、右陪集;当$gN=Ng$时,称其为N的一个**陪集(coset)**.

群G中包含一个正规子群$C\subset G$,称为群G的**中心(center)**,满足$\forall c\in C$与群G中的每个元g可交换,即$\forall c\in C,\forall g\in G\Rightarrow cg=gc$.

定理 1.2.11 设N为群G的子群,则N为正规子群,当且仅当$aNa^{-1}=N,\forall a\in G$.

证 必要性 若N是正规子群,则由假设条件有$aN=Na,\forall a\in G$. 于是$aNa^{-1}=(aN)a^{-1}=(Na)a^{-1}=N(aa^{-1})=N$. 必要性得证.

充分性 若$aNa^{-1}=N,\forall a\in G$,则$Na=(aNa^{-1})a=(aN)(a^{-1}a)=(aN)e=aN$,故$N$是正规子群.

群G关于正规子群N的陪集的个数,可能是有限的,也可能是无限的.

定义 1.2.17(正规子群的指数) 若群G关于正规子群N的陪集个数为有限数j,则称j为N在G内的**指数(index number)**.

定理 1.2.12 设G为有限群,$H\subset G$是G的正规子群,则H的阶n与它在G内的指数j都能整除G的阶N,并且满足$N=nj$. 进而,任一个元$a\in G$的阶都能整除G的阶.

证 设G的阶N、H的阶n、H在G内的指数j都是有限的正整数,G的N个元被H分成j个陪集,而且每个陪集中都有n个元,所以$N=nj$.

例 1.2.23 S_3中的子群$H=\{(1),(1\ \ 2)\}$不是S_3的正规子群,因为

$$(1\ \ 3)H=\{(1\ \ 3),(1\ \ 2\ \ 3)\},\quad H(1\ \ 3)=\{(1\ \ 3),(1\ \ 3\ \ 2)\}.$$

2. 积群

定义 1.2.18(积群、投影映射) 设$(G_1,(\cdot)_1),(G_2,(\cdot)_2)$是两个群,定义$G_1$与$G_2$的**积集(product set)**为

$$G_1\times G_2=\{(g_1,g_2):g_1\in G_1,g_2\in G_2\};$$

对于$(a_1,a_2)\in G_1\times G_2,(b_1,b_2)\in G_1\times G_2$,定义$G_1\times G_2$中的运算$\cdot$为:

$$(a_1,a_2)(b_1,b_2)\equiv(a_1,a_2)\cdot(b_1,b_2)=(a_1(\cdot)_1b_1,a_2(\cdot)_2b_2)\equiv(a_1b_1,a_2b_2).$$

不难验证,$(G_1\times G_2,\cdot)$构成一个新的群,称为G_1与G_2的**积群(product group)**.

积群$G=G_1\times G_2$与群G_1,G_2之间有两个重要的映射:**投影映射(projective mapping)** $\mathrm{Pr}_1:G\to G_1,\mathrm{Pr}_2:G\to G_2$,满足

$$\mathrm{Pr}_1(g_1,g_2)=g_1,\quad(g_1,g_2)\in G;$$

$$\mathrm{Pr}_2(g_1,g_2)=g_2,\quad(g_1,g_2)\in G.$$

它们分别是G到G_1与G_2的同态.

3. 商群

定义 1.2.19(等价类、商集)　设$(G,(\cdot)_G)$是一个群,$N\subset G$是G的正规子群. 对任意两个元$a,b\in G$,若$a^{-1}b\in N$,则称b与a(关于N)**等价(equivalent)**,记为$b\sim a$. 易知,$a^{-1}b\in N$当且仅当$b\in aN$.

记与$a\in G$等价的元$b\in G$的全体为$[a]$,亦即$[a]=\{b\in G:b\sim a\}\equiv aN$. 称$[a]$为$a$的**等价类(equivalent class)**. 也称$[a]$为a的**陪集(coset)**. 于是,G中的元得到一种新的组合,就是按等价关系$\sim$构成许多等价类. G中元的等价类的全体记为$\{[a]:a\in G\}=\{aN:a\in G\}$,并记$G/\sim\equiv G/N=\{[a]:a\in G\}$. 称$G/N$为$G$关于等价关系$\sim$的**商集(quotient set)**.

定理 1.2.13　设$N\subset G$是G的正规子群,$G/N\equiv\{[a]:a\in G\}$为商集,则$\sim$是一个等价关系:

(1) $a\sim a$,即$a\in[a]$;　(自反性)

(2) $a\sim b\Leftrightarrow b\sim a$,即$b\in[a]\Leftrightarrow a\in[b]$;　(对称性)

(3) $a\sim b,b\sim c\Rightarrow a\sim c$,即$b\in[a],c\in[b]\Rightarrow c\in[a]$;　(传递性)

进而,还有

(4) $a\in[b]\Rightarrow[a]=[b]$,即$a\sim b\Rightarrow[a]=[b]$;

(5) $[a]\neq[b]\Rightarrow[a]\cap[b]=\varnothing$,即$[a]\cap[b]\neq\varnothing\Rightarrow[a]=[b]$.

证明留作习题.

定义 1.2.20(商群、商映射)　在商集$G/N\equiv\{[a]:a\in G\}$中定义运算$\cdot$,$[a][b]\equiv[a]\cdot[b]=[a(\cdot)_G b]\equiv[ab]$,则等价类所成的商集$G/N$在此运算之下构成一个群,称为群$G$关于正规子群$N$的**商群(quotient group)**,记为

$$(G/N\equiv\{[a]:a\in G\},[a]\cdot[b]=[ab]).$$

定义映射$\pi:a\to\pi(a)=[a]$,称π为G到G/N上的**商映射(quotient mapping)**.

请读者验证:商映射是G到G/N上的同态,且是G到G/N上的满射.

例 1.2.24(剩余类群)　设$(\mathbb{Z},+)$为整数加群,取正整数$m\geqslant 2$,将$\mathbb{Z}$中所有m的倍数所成的集合记为$N=\{mk:k\in\mathbb{Z}\}$,则$(N,+)$是一个正规子群(请读者验证).

可以利用正规子群N将整数加群$(\mathbb{Z},+)$分为等价类. 定义等价类

$$[k]=\{mk:k\in\mathbb{Z}\}=kN,$$

则$\mathbb{Z}_m=\{[k]:k\in\mathbb{Z}\}$是关于正规子群$N$的等价类的全体,成为商集

$$\mathbb{Z}_m=\mathbb{Z}/N=\{[0],[1],\cdots,[m-1]\}\equiv\{0,1,\cdots,m-1\},$$

其中含有m个元素,在数的模m加法运算$x\oplus y=x+y(\mathrm{mod}\ m)$之下构成一个群,称为**剩余类群(residue class group)**.

例 1.2.25 置换群中的右(左)陪集.设

$$S_3=\{(1),(1\ \ 2),(1\ \ 3),(2\ \ 3),(1\ \ 2\ \ 3),(1\ \ 3\ \ 2)\}$$

为3-置换群,取子群 $H=\{(1),(1\ \ 2)\}$,例1.2.23告诉我们,H 不是正规子群,所以只能作右、左陪集.

做右陪集得

$$H(1)=\{(1),(1\ \ 2)\},$$
$$H(1\ \ 3)=\{(1\ \ 3),(1\ \ 3\ \ 2)\},$$
$$H(2\ \ 3)=\{(2\ \ 3),(1\ \ 2\ \ 3)\},$$
$$H(1\ \ 2)=\{(1\ \ 2),(1)\},$$
$$H(1\ \ 3\ \ 2)=\{(1\ \ 3\ \ 2),(1\ \ 3)\},$$
$$H(1\ \ 2\ \ 3)=\{(1\ \ 2\ \ 3),(2\ \ 3)\},$$

做左陪集得

$$(1)H=\{(1),(1\ \ 2)\},$$
$$(1\ \ 3)H=\{(1\ \ 3),(1\ \ 2\ \ 3)\},$$
$$(2\ \ 3)H=\{(2\ \ 3),(1\ \ 3\ \ 2)\},$$
$$(1\ \ 2)H=\{(1\ \ 2),(1)\},$$
$$(1\ \ 2\ \ 3)H=\{(1\ \ 2\ \ 3),(1\ \ 3)\},$$
$$(1\ \ 3\ \ 2)H=\{(1\ \ 3\ \ 2),(2\ \ 3)\}.$$

但是,$(1\ \ 3)H\neq H(1\ \ 3)$,$(2\ \ 3)H\neq H(2\ \ 3)$,$(1\ \ 3\ \ 2)H\neq H(1\ \ 3\ \ 2)$,$(1\ \ 2\ \ 3)H\neq H(1\ \ 2\ \ 3)$;只有 $(1)H=H(1)$,$(1\ \ 2)H=H(2\ \ 1)$,这是因为 H 只是一个子群而非正规子群.

定理 1.2.14 群 G 与它的每一个商群 G/N 同态.

证 设 G 是一个群,N 是它的任一正规子群;在 G 与商群 G/N 之间的商映射 $\pi:a\to[a](a\in G)$,即 $a\to aN(a\in G)$ 是 G 到 $G/N\equiv\{[a]:a\in G\}$ 的一个满射.进而,对于 $\forall\, a,b\in G$,有 $ab\to abN=(aN)(bN)$,所以 $a\to aN(a\in G)$ 是一个满同态,因此,群 G 与商群 G/N 同态.

习题 1

1. 试证集合的运算律(1)~(6).

2. 试证有理数集$\mathbb{Q}$的势是$\aleph_0$.

3. 设 X,Y 为两个基本集合,A,B,C 分别为它们的子集.设 $f:A\to B$ 是集合 A 到 B 的映射,集合 $f(A)=\{y\in B:y=f(x),x\in A\}$ 与 $f^{-1}(B)=\{x\in A:f(x)\in B\}$ 分别为集合 A 的像集与集合 B 的逆像集(或称原像集).试证:

(1) 若 $A\subset B$,则 $f(A)\subset f(B)$,若 $C\subset D$,则 $f^{-1}(C)\subset f^{-1}(D)$;

(2) $f(A\cup B)=f(A)\cup f(B)$，一般地，$f\left(\bigcup_{\lambda\in\Lambda}A_\lambda\right)=\bigcup_{\lambda\in\Lambda}f(A_\lambda)$，$\Lambda$ 为指标集，

$f(A\cap B)\subset f(A)\cap f(B)$，若 f 为单射，则成立等式，

$f(A\backslash B)\supset f(A)\backslash f(B)$，若 f 为单射，则等式成立；

(3) $f^{-1}(C\cup D)=f^{-1}(C)\cup f^{-1}(D)$，一般地，$f^{-1}\left(\bigcup_{\lambda\in\Lambda}C_\lambda\right)=\bigcup_{\lambda\in\Lambda}f^{-1}(C_\lambda)$，$\Lambda$ 为指标集，

$f^{-1}(C\cap D)=f^{-1}(C)\cap f^{-1}(D)$，一般地，$f^{-1}\left(\bigcap_{\lambda\in\Lambda}C_\lambda\right)=\bigcap_{\lambda\in\Lambda}f^{-1}(C_\lambda)$，$\Lambda$ 为指标集，

$f^{-1}(C\backslash D)=f^{-1}(C)\backslash f^{-1}(D)$；

(4) $X\backslash\left(\bigcup_{\lambda\in\Lambda}A_\lambda\right)=\bigcap_{\lambda\in\Lambda}(X\backslash A_\lambda)$，$X\backslash\left(\bigcap_{\lambda\in\Lambda}A_\lambda\right)=\bigcup_{\lambda\in\Lambda}(X\backslash A_\lambda)$，$\Lambda$ 为指标集；

(5) $\left(\bigcup_{\lambda\in\Lambda}A_\lambda\right)\cap\left(\bigcup_{\mu\in M}B_\mu\right)=\bigcup_{(\lambda,\mu)\in\Lambda\times M}(A_\lambda\cap B_\mu)$，

$\left(\bigcap_{\lambda\in\Lambda}A_\lambda\right)\cup\left(\bigcap_{\mu\in M}B_\mu\right)=\bigcap_{(\lambda,\mu)\in\Lambda\times M}(A_\lambda\cap B_\mu)$，$\Lambda,M$ 为指标集；

(6) $f^{-1}(f(A))\supset A$，$f(f^{-1}(C))\subset C$；

(7) 若 $f:A\to B$ 是单射，则 $f^{-1}(f(A))=A$；

(8) 若 $f:A\to B$ 是满射，则 $f(f^{-1}(B))=B$.

4. 试证：若群 G 的每个元 $x\in G$ 都适合 $x^2=e$，则 G 是一个交换群.

5. 一个有限群的每个元的阶都是有限数.

6. 设 $G=\mathbb{R}$，并设 G 的所有一一变换所成的集合为 S，并记为 $\tau:a\to a'=\tau(a)$，试证：定义 S 的乘法为复合：$\tau_1,\tau_2\in S\Rightarrow\tau_1[\tau_2(a)]=(\tau_1\circ\tau_2)(a)$，$a\in G$，则 S 构成一个群.

7. 找出 S_3 中不能与$\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix}$交换的元.

8. 证明定理 1.2.8、定理 1.2.9、定理 1.2.10.（在定理 1.2.9 中，嵌入映射 I 是子群 S 到像集 $I(S)$上的同构映射.）

9. 若 G 是循环群，并且 G 与$\widetilde{G}$同态，试证：$\widetilde{G}$也是循环群.

10. G 是无限阶循环群，$\widetilde{G}$是任一个循环群，试证：$\widetilde{G}$与 G 同态.

11. 试证：对于正整数 $m\geqslant 2$，集合 $N=\{mk:k\in\mathbb{Z}\}$在通常加法运算下成为正规子群.

12. 群$(G,\cdot)$有两个子群 G_1,G_2，定义群$(G,\cdot)$的直积 $G_1\times G_2$ 为

① 要求 G_1,G_2 都是正规子群；

② $G=G_1\times G_2=\{g_1\cdot g_2:g_1\in G_1,g_2\in G_2\}$，$G_1\cap G_2=\{e\}$，

其中$\{e\}$是 G 的单位元所成的子群．试证：群$(G,\cdot)$是它的两个子群 G_1,G_2 的直积的充分必要条件为

(1) G 中任一元 $g\in G$ 能惟一地表示为 $g=g_1g_2$，其中 $g_1\in G_1,g_2\in G_2$；

(2) G_1 中的元与 G_2 中的元可交换：$g_1g_2=g_2g_1$，其中 $g_1\in G_1,g_2\in G_2$.

第2章 线性空间与线性变换

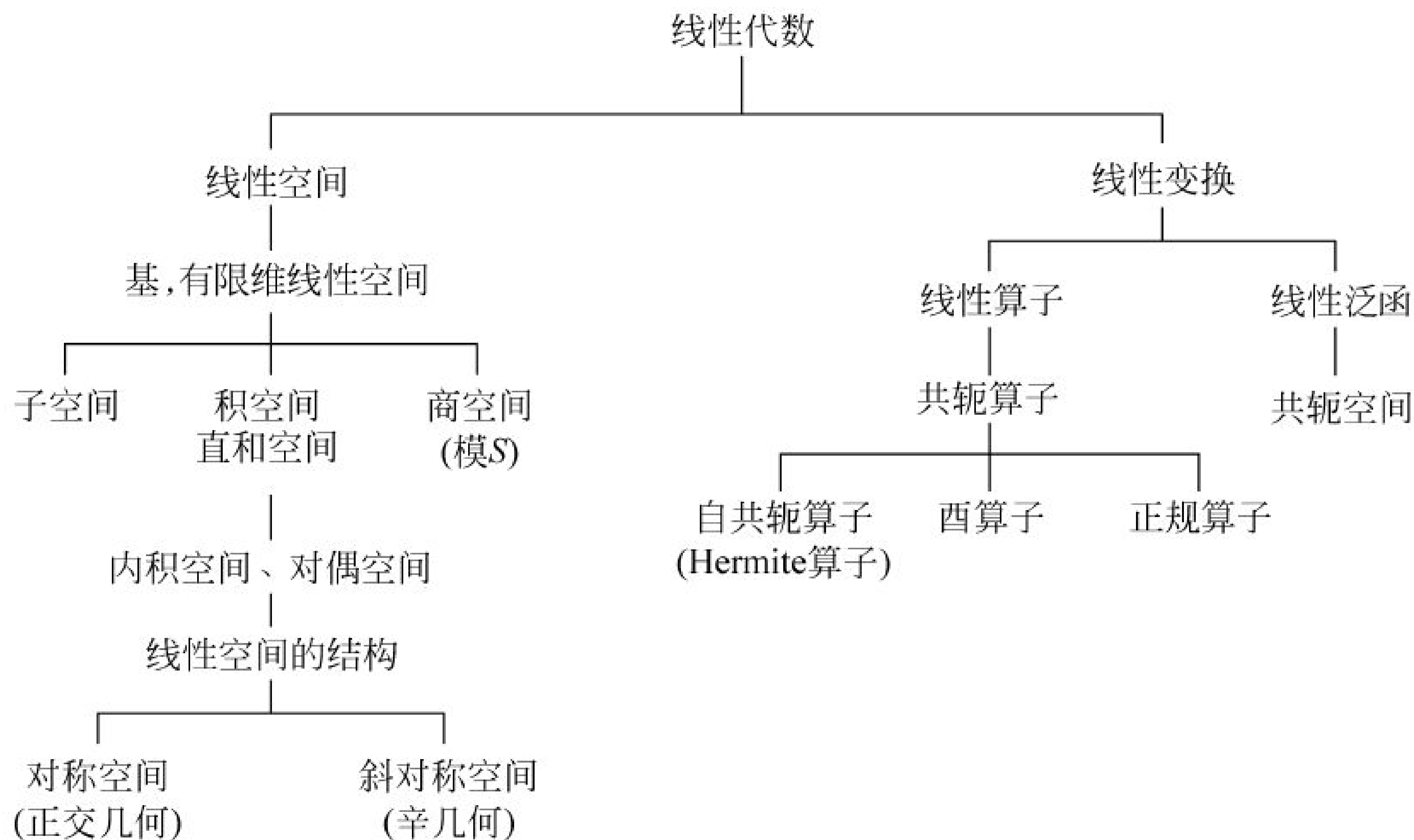
代数学，是数学诸分支中最早被人类认识并研究的一支．从算术开始，到用符号代替数字，进入“数字符号化”的时代，并逐渐转入解代数方程与代数方程组，形成“方程式论”的研究．19 世纪中叶，代数学从“方程式论”转向“代数运算”的研究，并得到数学家的一致认可，使得代数学与代数运算的一般理论与 20 世纪的近代观点统一起来．

近世代数的主要研究内容是集合以及在集合上的代数运算，使得集合对于运算而言形成“运算结构”，或“代数结构”，例如群、环、域、理想、模等．研究运算结构的性质与运算结构的表示，以及一种运算结构在另一种运算结构上的作用，成为近世代数的主要内容与任务．

线性代数则是研究线性空间、模（线性空间的推广）、作用在线性空间或模之间的线性变换（也称为线性映射、线性算子），以及相关问题的数学科学分支．本节以线性空间的结构与其间的线性变换为主来介绍线性代数的基本知识．

高等数学中关于线性代数的知识需要熟练掌握，用以作为学习本章的例子与基础．本章主要参考文献为[11]，[13]．

关于本章的内容，其教学框图如下：



2.1 线性空间

2.1.1 线性空间的实例

线性空间的定义已在1.2节由定义1.2.5给出，并且有两个例子：n维欧氏空间$\mathbb{R}^n$与区间$[a,b]$上的连续函数空间$C([a,b])$. 本节开始，再给出一些重要的实例.

1. 矩阵空间

$m\times n$实值矩阵的全体$\mathfrak{M}_{m\times n}(\mathbb{R})=\{[a_{ij}]:1\leqslant i\leqslant m,1\leqslant j\leqslant n,a_{ij}\in\mathbb{R}\}$，在矩阵加法$+$与数乘$\alpha\cdot$之下，成为实数域$\mathbb{R}$上的线性空间，即

$$A=[a_{ij}],\ B=[b_{ij}]\quad\Rightarrow\quad A+B=[a_{ij}+b_{ij}];$$

$$A=[a_{ij}],\ \alpha\in\mathbb{R}\quad\Rightarrow\quad \alpha\cdot A=[\alpha\, a_{ij}].$$

称$\mathfrak{M}_{m\times n}(\mathbb{R})$为**实数域$\mathbb{R}$上的矩阵空间(matrix space)**；若$a_{ij}\in\mathbb{C}$，$\alpha\in\mathbb{C}$，则称$\mathfrak{M}_{m\times n}(\mathbb{C})$为**复数域$\mathbb{C}$上的矩阵空间.**

记$\mathfrak{M}_n\equiv\mathfrak{M}_{n\times n}(\mathbb{R})=\{A=[a_{ij}]:a_{ij}\in\mathbb{R}\}$，其中$A=[a_{ij}]=\begin{bmatrix}a_{11}&\cdots&a_{1n}\\\vdots&\ddots&\vdots\\a_{n1}&\cdots&a_{nn}\end{bmatrix}$称为$n$阶方阵，$\mathfrak{M}_n$称为**实方阵空间(real square matrix space)**；相应地，可定义**复方阵空间.**

2. l^2空间

设$l^2=\left\{x:x=(x_1,\cdots,x_n\cdots),x_j\in\mathbb{R},j=1,2,\cdots,n,\cdots,\sum\limits_{j=1}^{+\infty}|x_j|^2<+\infty\right\}$，定义其加法运算$+$与数乘运算$\alpha\cdot$为

$$x+y=(x_1,x_2,\cdots)+(y_1,y_2,\cdots)=(x_1+y_1,x_2+y_2,\cdots),$$

$$\alpha x\equiv\alpha\cdot x=\alpha\cdot(x_1,x_2,\cdots)=(\alpha x_1,\alpha x_2,\cdots),\quad \alpha\in\mathbb{R},$$

则l^2在这种运算结构下成为一个线性空间. 显见，l^2空间是欧氏空间$\mathbb{R}^n$的自然推广.

3. 数域$\mathbb{F}$上的多项式空间$\mathbb{F}[x]$

以数域$\mathbb{F}$的元为系数的一元多项式

$$p(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0,\quad a_j\in\mathbb{F},\ j=0,1,2,\cdots,n,$$

其全体记为

$$\mathbb{F}[x]=\{p:p(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0,\quad a_j\in\mathbb{F},\ n\in\mathbb{N},\ j=0,1,2,\cdots,n\}.$$

在$\mathbb{F}[x]$上定义两个元$p,q\in\mathbb{F}[x]$的加法$+$与数乘$\alpha\cdot$为

$$(p+q)(x)=p(x)+q(x)$$
$$=(a_n+b_n)x^n+(a_{n-1}+b_{n-1})x^{n-1}+\cdots+(a_1+b_1)x+(a_0+b_0),$$
$$(\alpha\cdot p)(x)\equiv \alpha p(x)=\alpha a_n x^n+\alpha a_{n-1}x^{n-1}+\cdots+\alpha a_1 x+\alpha a_0,\quad \alpha\in\mathbb{F}.$$

不难验证,$\mathbb{F}[x]$成为数域$\mathbb{F}$上的线性空间,称为**数域 $\mathbb{F}$ 上的多项式空间(polynomial space)**.

4. L 可积函数空间 $L^1(E)$,L 平方可积函数空间 $L^2(E)$

称 L 可测集 E 上 L 可积函数的全体为 L 可积函数类,记为

$$L^1(E)=\left\{f:\int_E |f(x)|\,\mathrm{d}m<+\infty\right\};$$

称集合 E 上 L 平方可积函数的全体为 L 平方可积函数类,记为

$$L^2(E)=\left\{f:\int_E |f(x)|^2\,\mathrm{d}m<+\infty\right\}.$$

定义其加法运算+与数乘运算 $\alpha\cdot$为

$$(f+g)(x)=f(x)+g(x),\quad x\in E;\quad (\alpha\cdot f)(x)=\alpha\cdot f(x),\quad x\in E,\alpha\in\mathbb{F}.$$

在这种运算结构之下,二者都成为线性空间,$L^1(E)$与 $L^2(E)$分别称为 **L 可积函数空间(Lebesgue integrable function space)**与 **L 平方可积函数空间(Lebesgue square integrable function space)**.

2.1.2 线性空间的基

1. 几个基本概念

1) 线性无关子集

设$(X,+,\alpha\cdot)$是数域$\mathbb{F}$上的线性空间,$S\subset X$ 是 X 的非空子集. 若 S 中的任意 n 个元 $a_j\in S(j=1,2,\cdots,n)$都是线性无关的(亦即,对$\mathbb{F}$中的任意 n 个数 $r_j\in\mathbb{F}$ $(j=1,2,\cdots,n)$,当 $r_1a_1+r_2a_2+\cdots+r_na_n=0$ 时,总有 $r_1=r_2=\cdots=r_n=0$),则称 S 为 X 的**线性无关子集(linear independent subset)**.

若 X 的非空子集 S 不是线性无关的,则称 S 为**线性相关(linear dependent)**的.

2) 由子集 $S\subset X$ 生成 X

若 $S\subset X$ 是数域$\mathbb{F}$上线性空间 X 的非空子集,又若 X 中的任一元 $x\in X$ 都可表示成 S 中的有限个元的线性组合,亦即,$\forall x\in X$,存在 $a_j\in S$ 与 $r_j\in\mathbb{F}$ $(j=1,2,\cdots,m)$,使得 $x=r_1a_1+r_2a_2+\cdots+r_ma_m$,则称**子集 S 生成(generate)**X. 记为 $X=\mathrm{span}(S)$. 不失一般性,我们假设 X 的生成集 S 是线性无关子集.

非空线性无关子集 $S\subset X$ 的元的有限线性组合组成 X 中的一个线性子空间,记为$\langle S\rangle$,即

$$\langle S\rangle=\mathrm{span}(S)=\left\{a=\sum_{j=1}^{m}r_ja_j:a_j\in S,r_j\in\mathbb{F},j=1,2,\cdots,m,m\in\mathbb{Z}^+\right\}.$$

于是有$\varnothing\neq S\subset X\Rightarrow\langle S\rangle\subset X$,且$\langle S\rangle$是$\mathbb{F}$上的一个线性空间(请读者自行验证).

3）线性空间之间的同构

设 X_1,X_2 是数域$\mathbb{F}$上的两个线性空间，若存在双方单值的满射 $T:X_1\to X_2$，使得

$$T(x_1+x_2)=T(x_1)+T(x_2),\quad x_1,x_2\in X_1;$$

$$T(\alpha x_1)=\alpha T(x_1),\quad x_1\in X_1,\quad \alpha\in\mathbb{F},$$

则称 T 为 X_1 到 X_2 的同构映射，简称同构；并称 X_1 与 X_2 为**同构空间**，简称 X_1 与 X_2 **同构**.

注 同构映射是一对一的满映射，且保持线性空间的运算. 请读者回顾 1.2 节关于两个群之间同构映射的定义 1.2.8，并将这两种同构进行比较.

2. 线性空间的基

定义 2.1.1（线性空间的基） 设$(X,+,\alpha\cdot)$为数域$\mathbb{F}$上的线性空间，若线性无关的子集 $S\subset X$ 生成 X，即 $X=\mathrm{span}(S)$，则称 S 为 X 的一个**基**(**basis**).

定理 2.1.1 对任何线性空间 X，若 $X\neq\{0\}$，则 X 一定存在基.

证 设线性空间 $X\neq\{0\}$，将 X 中的线性无关子集 $S\subset X$ 的全体记为

$$\mathfrak{B}=\{S\subset X: S\neq\varnothing \text{ 为 } X \text{ 的线性无关子集}\}.$$

显然 $\mathfrak{B}\neq\varnothing$，因为 $\forall S=\{a\in X: a\neq 0\}\in\mathfrak{B}$是仅含一个非零元的线性无关的子集.

对于集合 $\mathfrak{B}$将其中的元素按照集合的包含关系定义一个“序”：$S_1\subset S_2 \Leftrightarrow S_1$ 包含于 S_2. 于是，$\mathfrak{B}$成为一个**半序集**(**partial order set**)（由于 $\mathfrak{B}$中的集合之间有的能按照包含关系排序，有的不能，因此其中只有一部分能排序，这样的 $\mathfrak{B}$称为一个半序集[6]）. 若任取 $\mathfrak{B}$中的一个全序的线性无关集合的“链”$S_1\subset S_2\subset\cdots\subset S_k\subset\cdots$（$S_k\in\mathfrak{B}$，则并集 $U=\bigcup\limits_{k\geqslant 1}S_k$ 仍然是 $\mathfrak{B}$中的线性无关的子集，U 也是链 $S_1\subset S_2\subset\cdots\subset S_k\subset\cdots$的上界. 于是，由 Zorn 引理（见[6]：若集合 P 是一个半序集，且其中每个链都有上界，则 P 有极大元）知，$\mathfrak{B}$中必有极大元$\widetilde{S}\in\mathfrak{B}$亦即，$\widetilde{S}$中的元是线性无关的，并且所有包含$\widetilde{S}$的子集 $V\subset X(\widetilde{S}\subsetneqq V)$都不再是线性无关的.

于是，$\widetilde{S}$一定生成 X. 因若不然，必有元 $x\in X\backslash\widetilde{S}$，使得 x 不是$\widetilde{S}$的元的线性组合，而且集合$\widetilde{S}\cup\{x\}$是 $\mathfrak{B}$中的线性无关的子集，$\widetilde{S}$又是$\widetilde{S}\cup\{x\}$的真子集，矛盾. 定理得证.

定义 2.1.2（线性空间的维数） 设$(X,+,\alpha\cdot)$为域$\mathbb{F}$上的线性空间，S 为 X 的基. 称 S 的基数$\overline{\overline{S}}=\mathrm{card}(S)$为 X 的**维数**(**dimension**)，记为 $\dim X=\mathrm{card}(S)$.

若 $\mathrm{card}(S)$为有限数 $n(n\in\mathbb{Z}^+)$，则称 X 为（有限）n **维线性空间**. 例如$(\mathbb{R}^n,+,\alpha\cdot)$，它的一组标准正交基是

$$(\underbrace{1,0,0,\cdots,0}_{n}),(\underbrace{0,1,0,\cdots,0}_{n}),\cdots,(\underbrace{0,0,\cdots,0,1}_{n}),$$

含 n 个线性无关的向量，因此$\mathbb{R}^n$ 是有限维线性空间，其维数为 n.

本书中约定：有限 n 维线性空间简称 n **维线性空间**，不再加“有限”两字.

在线性代数领域中，主要研究 n 维线性空间.

命题 2.1.1 n 维线性空间维数的定义 2.1.2 是合理的、准确的.

证 ① 若数域$\mathbb{F}$上的线性空间$(X,+,\alpha\cdot)$中的一组元 $V=\{v_1,v_2,\cdots,v_m\}$线性无关；另一组元 $S=\{s_1,s_2,\cdots,s_n\}(n=\text{card}(S))$生成 X,则 $m\leqslant n$.

事实上,由 $S=\{s_1,s_2,\cdots,s_n\}$生成 X,知 $\forall x\in X\Rightarrow x=\sum\limits_{j=1}^{n}\alpha_j s_j$, $\alpha_j\in\mathbb{F}$, $j=1,2,\cdots,n$. 列出元组 $s_1,\cdots,s_n$; $v_1,\cdots,v_m$. 将最后一个元移至第一个,得到 $v_m,s_1,\cdots,s_n$; $v_1,\cdots,v_{m-1}$.

由于 $S=\{s_1,s_2,\cdots,s_n\}$生成 X,故 v_m 可表示成 $s_1,s_2,\cdots,s_n$ 的线性组合,因此可从 $s_1,s_2,\cdots,s_n$ 中移走一个,例如移走 s_j,使得$\{v_m,s_1,\cdots,s_{j-1},s_{j+1},\cdots,s_n\}$仍生成线性空间 X,记为 $\{v_m,s_1,\cdots,\hat{s}_j,\cdots,s_n\}$,$\hat{s}_j$ 表示 s_j 已被移走. 于是得到一个新向量集$\{v_m,s_1,\cdots,\hat{s}_j,\cdots,s_n;\ v_1,\cdots,v_{m-1}\}$.

重复以上过程,得到另一个新的向量组$\{v_{m-1},v_m,s_1,\cdots,\hat{s}_j,\cdots,s_n;\ v_1,\cdots,v_{m-2}\}$. 同理,可从前一个集合中移走某个 s_k,使得移走后的一个集合仍可生成 X,从而得到

$$\{v_{m-1},v_m,s_1,\cdots,\hat{s}_j,\cdots,\hat{s}_k,\cdots,s_n;\ v_1,\cdots,v_{m-2}\}.$$

如此一直进行下去,我们称此移走的方法为对集合$\{s_1,s_2,\cdots,s_m\}$进行 Steinitz 替换. 替换的结果是:每次移完后,得到的新向量集都仍然可以生成 X.

若所有的 $s_1,s_2,\cdots,s_n$ 首先被移完,亦即 $n<m$,则集合$\{v_{m-n},v_{m-n+1},\cdots,v_{m-1},v_m,\hat{s}_1,\cdots,\hat{s}_n\}$成为 $V=\{v_1,v_2,\cdots,v_m\}$的真子集$\bar{V}_j\equiv\{v_{m-n},v_{m-n+1},\cdots,v_{m-1},v_m,\hat{s}_1,\cdots,\hat{s}_n\}\subset V$,而且又能生成线性空间 X,这与 $V=\{v_1,v_2,\cdots,v_m\}$线性无关相矛盾(因为 $v_{m-n-1},v_{m-n-2},\cdots,v_1\notin\bar{V}_j$,却可以用$\bar{V}_j(\subset V)$中的元线性表示,这说明 V 不是线性无关集),故必须有 $m\leqslant n$.

② 对换关于 S 与 V 的假设,设 $S=\{s_1,s_2,\cdots,s_n\}$线性无关,而 $V=\{v_1,v_2,\cdots,v_m\}(m=\text{card}(V))$是线性空间 X 的生成集,$X=\text{span}(V)$,则 $n\leqslant m$. 这只要用①的论证方法便可得到.

③ 当 $S=\{s_1,s_2,\cdots,s_n\}$与 $V=\{v_1,v_2,\cdots,v_m\}$都是 X 的生成集(因而也都是线性无关的)时,必有 $m=n$. 这是显然的.

由①～③立即得到,空间 X 的基的定义是合理的,准确的.

定理 2.1.2 设$(X,+,\alpha\cdot)$为数域$\mathbb{F}$上的 n 维线性空间,$S\subset X$ 为一子集,则下述论断等价:

(1) S 是 X 的基;

(2) X 中每个元 $x\in X$ 可惟一地表示为

$$x=r_1v_1+r_2v_2+\cdots+r_nv_n,\quad r_j\in\mathbb{F},\quad v_j\in S,\quad j=1,2,\cdots,n;$$

(3) S 是 X 中的极小生成集(X 的所有生成集都包含 S);

(4) S 是 X 中的极大线性无关集(X 的所有线性无关集都被 S 所包含).

特别指出,n 维线性空间的结构与它的数域有密切关系.

定理 2.1.3 设$(X,+,\alpha\cdot)$为数域$\mathbb{F}$上的n维线性空间,则

(1) $\mathbb{F}$上的n维线性空间X与$\mathbb{F}^n$同构;亦即,实数域$\mathbb{R}$上的n维线性空间X与$\mathbb{R}^n$同构、复数域$\mathbb{C}$上的n维线性空间X与$\mathbb{C}^n$同构;

(2) $\mathbb{F}$上两个n维线性空间X与Y同构,当且仅当它们的维数相同.

证 只需证明(1). 为证$\mathbb{F}$上的n维线性空间$(X,+,\alpha\cdot)$与$\mathbb{F}^n$同构,取X的一组基$\mathfrak{B}=\{b_1,b_2,\cdots,b_n\}$,则$\forall x\in X,\exists(\alpha_1,\alpha_2,\cdots,\alpha_n)\in\mathbb{F}^n$,使得$x=\alpha_1b_1+\alpha_2b_2+\cdots+\alpha_nb_n$,并且用列向量$[x]_{\mathfrak{B}}=\begin{bmatrix}\alpha_1\\\alpha_2\\\vdots\\\alpha_n\end{bmatrix}$表示,称$[x]_{\mathfrak{B}}$为$x\in X$在基$\mathfrak{B}$之下的**坐标(coodinate)**. 定义一个映射$\varphi_{\mathfrak{B}}:X\to\mathbb{F}^n$,其取值为

$$\varphi_{\mathfrak{B}}(x)=[x]_{\mathfrak{B}}\leftrightarrow(\alpha_1,\alpha_2,\cdots,\alpha_n)\in\mathbb{F}^n,\quad x\in X.$$

显然,$\varphi_{\mathfrak{B}}$是X到$\mathbb{F}^n$上的一对一的、线性的满映射. 进而,$\varphi_{\mathfrak{B}}$是X到$\mathbb{F}^n$的同构映射,它保持运算:

$$\varphi_{\mathfrak{B}}(x+y)=[x+y]_{\mathfrak{B}}=\begin{bmatrix}\alpha_1+\beta_1\\\alpha_2+\beta_2\\\vdots\\\alpha_n+\beta_n\end{bmatrix}_{\mathfrak{B}}=\begin{bmatrix}\alpha_1\\\alpha_2\\\vdots\\\alpha_n\end{bmatrix}_{\mathfrak{B}}+\begin{bmatrix}\beta_1\\\beta_2\\\vdots\\\beta_n\end{bmatrix}_{\mathfrak{B}}=\varphi_{\mathfrak{B}}(x)+\varphi_{\mathfrak{B}}(y),$$

$$\varphi_{\mathfrak{B}}(\alpha x)=[\alpha x]_{\mathfrak{B}}=\begin{bmatrix}\alpha\alpha_1\\\alpha\alpha_2\\\vdots\\\alpha\alpha_n\end{bmatrix}_{\mathfrak{B}}=\alpha\begin{bmatrix}\alpha_1\\\alpha_2\\\vdots\\\alpha_n\end{bmatrix}_{\mathfrak{B}}=\alpha\varphi_{\mathfrak{B}}(x).$$

于是,在**同构**的意义下,n维线性空间只有一个,就是$\mathbb{F}^n$.

例 2.1.1 n维欧氏空间$(\mathbb{R}^n,+,\alpha\cdot)$是$n$维线性空间,它有基

$$e_1=(1,0,0,\cdots,0),\quad e_2=(0,1,0,\cdots,0),\quad\cdots,\quad e_n=(0,0,0,\cdots,1).$$

例 2.1.2 $m\times n$实值矩阵的全体所组成的$\mathbb{R}$上的线性空间

$$\mathfrak{M}_{m\times n}=\{A=[a_{ij}]:1\leqslant i\leqslant m,1\leqslant j\leqslant n,a_{ij}\in\mathbb{R}\},$$

它有基

$$\mathfrak{B}_{m\times n}=\{[e_{ij}]:1\leqslant i\leqslant m,1\leqslant j\leqslant n\},$$

其中$[e_{ij}]=[a_{ks}]_{m\times n}$, $a_{ks}=\begin{cases}1, & k=i,s=j,\\0, & \text{其他}.\end{cases}$

易见,$\dim\mathfrak{M}_{m\times n}=mn$.

例 2.1.3 次数不超过n的多项式的全体(含零多项式)组成$\mathbb{R}$上的线性空间

$$\mathscr{P}_n=\{p_n(x):p_n(x)=a_nx^n+\cdots+a_1x+a_0,a_j\in\mathbb{R},j=0,1,\cdots,n\},$$

它有基$\mathfrak{B}=\{1,x,x^2,\cdots,x^n\}$. 其维数为$\dim\mathscr{P}_n=n+1$.

2.1.3 线性空间的子空间、积空间、直和空间、商空间

1. 子空间

设$(X,+,\alpha\cdot)$是数域$\mathbb{F}$上的线性空间，$S\subset X$是X的非空子集，如何由$(X,+,\alpha\cdot)\equiv(X,(+)_X,(\alpha\cdot)_X)$生成一个子线性空间$(S,+,\alpha\cdot)\equiv(S,(+)_S,(\alpha\cdot)_S)$呢？

子集　$S\subset X$是X的非空子集，$S\neq\varnothing$.

运算　$(+)_S$　$x,y\in S\Rightarrow x\,(+)_S y=x\,(+)_X y$；

$(\alpha\cdot)_S$　$x\in S,\ \alpha\in\mathbb{F}\Rightarrow(\alpha\cdot)_S x=(\alpha\cdot)_X x$.

若S在上述运算$(+)_S$，$(\alpha\cdot)_S$下成为数域$\mathbb{F}$上的线性空间，则称$(S,(+)_S,(\alpha\cdot)_S)$为$(X,(+)_X,(\alpha\cdot)_X)$的**线性子空间(linear subspace)**，简称**子空间(subspace)**.

为验证一个子集S是否构成X的子空间，只需验证线性空间的运算$(+)_X$与$(\alpha\cdot)_X$在子集S中是否封闭就够了.

嵌入映射　子空间S到线性空间X的映射$I:S\to X$，$I(x)=x$，$x\in S$，称为**嵌入映射**.

嵌入映射是S到$S\subset X$的同构映射，它是联系线性空间X与子空间S之间的重要桥梁.

定理 2.1.4　设$(X,+,\alpha\cdot)$为域$\mathbb{F}$上的n维线性空间，$S,T\subset X$是两个线性子空间，则$\dim(S)+\dim(T)=\dim(S+T)+\dim(S\cap T)$. 特别地，当$S\cap T=\{0\}$时，有

$$\dim(S)+\dim(T)=\dim(S+T)=\dim X=n.$$

2. 积空间

如何由数域$\mathbb{F}$上的两个线性空间$(X,(+)_X,(\alpha\cdot)_X)$，$(Y,(+)_Y,(\alpha\cdot)_Y)$生成一个积线性空间$(X\times Y,+,\alpha\cdot)\equiv(X\times Y,(+)_{X\times Y},(\alpha\cdot)_{X\times Y})$呢？

积集　$X\times Y=\{(x,y):x\in X,y\in Y\}$.

运算$(+)_{X\times Y}$　$(x_1,y_1),(x_2,y_2)\in X\times Y\Rightarrow$

$$(x_1,y_1)(+)_{X\times Y}(x_2,y_2)=(x_1(+)_X x_2,y_1(+)_Y y_2);$$

$(\alpha\cdot)_{X\times Y}$　$(x,y)\in X\times Y,\alpha\in\mathbb{F}\Rightarrow(\alpha\cdot)_{X\times Y}(x,y)=((\alpha\cdot)_X x,(\alpha\cdot)_Y y)$.

不难证明，$X\times Y$在上述运算$(+)_{X\times Y}$，$(\alpha\cdot)_{X\times Y}$下成为数域$\mathbb{F}$上的线性空间，称

$$(X\times Y,+,\alpha\cdot)\equiv(X\times Y,(+)_{X\times Y},(\alpha\cdot)_{X\times Y})$$

为$(X,(+)_X,(\alpha\cdot)_X)$与$(Y,(+)_Y,(\alpha\cdot)_Y)$的**积线性空间**，简称**积空间(product space)**.

联系积空间$(X\times Y,(+)_{X\times Y},(\alpha\cdot)_{X\times Y})$与原空间$(X,(+)_X,(\alpha\cdot)_X)$，$(Y,(+)_Y,(\alpha\cdot)_Y)$的重要桥梁是投影映射.

投影映射　称映射$\mathrm{Pr}_1:X\times Y\to X$，$\mathrm{Pr}_1((x,y))=x$与映射$\mathrm{Pr}_2:X\times Y\to Y$，$\mathrm{Pr}_2((x,y))=y$分别为积空间$X\times Y$到空间$X$与$Y$的**投影映射**.

易证，Pr_1与Pr_2为多对一的同态映射.

两个空间的积空间概念可以推广到一般m个线性空间的积空间情形，即

$$(X_1\times\cdots\times X_m,+,\alpha\cdot)\equiv(X_1\times\cdots\times X_m,(+)_{X_1\times\cdots\times X_m},(\alpha\cdot)_{X_1\times\cdots\times X_m}).$$

相应地,投影映射也有 m 个,即

$$\mathrm{Pr}_j:X_1\times\cdots\times X_m\to X_j,\quad \mathrm{Pr}_j((x_1,\cdots,x_m))=x_j,\quad j=1,2,\cdots,m.$$

3. 直和空间

由数域$\mathbb{F}$上两个线性空间$(X,+,\alpha\cdot)$与$(Y,+,\alpha\cdot)$可生成另一种有用的线性空间——直和空间.

设$(X,+,\alpha\cdot)\equiv(X,(+)_X,(\alpha\cdot)_X)$与$(Y,+,\alpha\cdot)\equiv(Y,(+)_Y,(\alpha\cdot)_Y)$为数域$\mathbb{F}$上的两个线性空间,且 $X\cap Y=\{0\}$.

直和 $X\oplus Y=\{(x,y):x\in X,y\in Y\}$称为 X 与 Y 的**直和(direct sum)**.

运算$(+)_{X\oplus Y}$ $(x_1,y_1),(x_2,y_2)\in X\oplus Y\Rightarrow(x_1,y_1)(+)_{X\oplus Y}(x_2,y_2)=(x_1(+)_Xx_2,y_1(+)_Yy_2)$;

$(\alpha\cdot)_{X\oplus Y}$ $(x,y)\in X\oplus Y,\alpha\in\mathbb{F}\Rightarrow(\alpha\cdot)_{X\oplus Y}(x,y)=((\alpha\cdot)_Xx,(\alpha\cdot)_Yy)$.

记为$(X\oplus Y,+,\alpha\cdot)\equiv(X\oplus Y,(+)_{X\oplus Y},(\alpha\cdot)_{X\oplus Y})$,不难验证,$(X\oplus Y,(+)_{X\oplus Y},(\alpha\cdot)_{X\oplus Y})$构成$\mathbb{F}$上的一个线性空间,称为$(X,+,\alpha\cdot)$与$(Y,+,\alpha\cdot)$的**直和空间(direct sum space)**,简称**直和**.

两个空间的直和概念也可以推广到一般 m 个线性空间的直和空间的情形.

$$(X_1\oplus\cdots\oplus X_m,+,\alpha\cdot)\equiv(X_1\oplus\cdots\oplus X_m,(+)_{X_1\oplus\cdots\oplus X_m},(\alpha\cdot)_{X_1\oplus\cdots\oplus X_m}).$$

由直和概念可以引进重要的"补集"概念.

定义 2.1.3(补集) 若 $S,T\subset X$ 是数域$\mathbb{F}$上的线性空间 X 的两个子空间,且 $X=S\oplus T,S\cap T=\{0\}$,则称 S 与 T **互为补集(complementary set)**,简称**互补**,记为 $S=\mathfrak{C}T$ 与$T=\mathfrak{C}S$.

定理 2.1.5 设$(X,+,\alpha\cdot)$为数域$\mathbb{F}$上的 n 维线性空间,$S\subset X$ 是一个线性子空间,则 $\dim X=\dim(S)+\dim(\mathfrak{C}S)$.

例 2.1.4 设$(\mathbb{R}^n,+,\alpha\cdot)$,且$(\mathbb{R}^m,+,\alpha\cdot)$,$m<n$ 是$\mathbb{R}^n$ 的 m 维线性子空间. 则$(\mathbb{R}^{n-m},+,\alpha\cdot)$也是$\mathbb{R}^n$ 的线性子空间,且有 $\mathbb{R}^n=\mathbb{R}^m\oplus\mathbb{R}^{n-m}$.

例 2.1.5 在方阵空间 $\mathfrak{M}_n$中,定义子集 $W=\left\{[a_{ij}]_{n\times n}\in\mathfrak{M}_n:\sum_{j=1}^n a_{1j}=0\right\}$,则 W 是 $\mathfrak{M}_{\times n}$的一个线性子空间.

显然 W 关于 $\mathfrak{M}_n$的加法与数乘运算是封闭的,且零方阵属于 W,即 $O=[0]_{n\times n}\in W$,因此 W 成为 $\mathfrak{M}_{\times n}$的线性子空间. 由于 W 中第一行的元素中只有 $n-1$ 个是线性无关的,故 $\dim W=n^2-1$.

4. 商空间

由数域$\mathbb{F}$上的线性空间$(X,+,\alpha\cdot)$与其上的一个等价关系$\sim$,可生成一个商线性空间

$$(X/\sim,+,\alpha\cdot)\equiv(X/\sim,(+)_{X/\sim},(\alpha\cdot)_{X/\sim}).$$

等价关系～ $x,y\in X$,称 x 与 y **等价**,记为 $x\sim y$,若满足

(1) $x\sim x$; (自反性 reflexivity)

(2) $x\sim y\Rightarrow y\sim x$; (对称性 symmetry)

(3) $x\sim y,y\sim z\Rightarrow x\sim z$. (传递性 transitivity)

等价类$[x]$ 集合$[x]=\{y\in X:y\sim x\}$称为 $x\in X$ 的**等价类**,也称$[x]$为 x 的**陪集**.

商集 $X/\sim=\{[x]:x\in X\}$称为线性空间 X 在等价关系～下的**商集**.

运算 $(+)_{X/\sim}$ $[x],[y]\in X/\sim\Rightarrow[x](+)_{X/\sim}[y]=[x\,(+)_X y]$;

$(\alpha\cdot)_{X/\sim}$ $[x]\in X/\sim,\ \alpha\in\mathbb{F}\Rightarrow(\alpha\cdot)_{X/\sim}([x])=[(\alpha\cdot)_X x]$.

易证,$X/\sim$在上述运算$(+)_{X/\sim}$,$(\alpha\cdot)_{X/\sim}$之下成为数域$\mathbb{F}$上的线性空间,称线性空间

$$(X/\sim,+,\alpha\cdot)\equiv(X/\sim,(+)_{X/\sim},(\alpha\cdot)_{X/\sim})$$

为$(X,+,\alpha\cdot)$在等价关系～下的**商线性空间**,简称**商空间**(**quotient space**).

商映射 称线性空间 X 到线性空间 $X/\sim$的映射 $\pi:x\to\pi(x)=[x]$为**商映射**. π 是多对一的同态映射.

例 2.1.6 在三维欧氏空间$\mathbb{R}^3$ 与其子空间$\mathbb{R}^2$ 之间,定义一个映射 $T:\mathbb{R}^3\to\mathbb{R}^2$,使得 $\forall x=(x_1,x_2,x_3)\in\mathbb{R}^3,T(x)=(x_1,x_2)\in\mathbb{R}^2$. 由此映射,定义

$$\ker(T)=\{(0,0,x_3)\in\mathbb{R}^3:x_3\in\mathbb{R}\},$$

称集合 $\ker(T)$为**映射 T 的核**(**kernel**);也定义

$$\mathrm{im}(T)=T(\mathbb{R}^3)=\mathbb{R}^2,$$

称集合 $\mathrm{im}(T)$为**映射 T 的像**(**image**).

考虑等价关系～:$x,y\in\mathbb{R}^3\Rightarrow x-y\in\ker(T)$. 不难验证,商集为

$$\mathbb{R}^3/\sim=\mathrm{im}(T),$$

这个商集在等价类的运算下构成的商空间就是$\mathbb{R}^2$.

5. 模 S 同余

作为商空间概念的推广,我们引进模 S 同余与模 S 商空间.

定义 2.1.4(**模 S 同余**) 设$(X,+,\alpha\cdot)$为域 F 上的线性空间,$S\subset X$ 为子空间,若

$$u,v\in X\Rightarrow u-v\in S,$$

则称 u 与 v **模 S 同余**(**congruent** (**modulo** S)),记为 $u\equiv v(\mathrm{mod}\ S)$.

与 $v\in X$ 模 S 同余的全体记为$[v]=\{u\in X:u\equiv v(\mathrm{mod}\ S)\}$,称为 v 的模 S 同余类,或称$[v]$为线性空间 X 中 v 的模 S 陪集.

模 S 同余是一个等价关系～,称为模 S 等价关系,它将X 划分成若干个“块”$[v]$.

6. 模 S 商空间

定义 2.1.5（模 S 商空间） 设$(X,+,\alpha\cdot)$为域 F 上的线性空间，$S\subset X$ 为子空间，$\sim$为 X 中的模 S 等价关系，模 S 同余类的全体记为

$$X/\sim\equiv X/S=\{[v]:v\in X\}=\{[v]=v+S:v\in X\},$$

并在其上定义加法与域 F 的乘法为

$$[u]+[v]=(u+S)+(v+S)=(u+v)+S,\quad u,v\in X;$$
$$\alpha[u]=\alpha(u+S)=\alpha u+S,\quad u\in X,\ \alpha\in F,$$

得到的 F 上的线性空间 $X/\sim=X/S$，称为 X 的**模 S 商空间(quotient space of X modulo S)**.

2.1.4 内积空间

最重要、最常用的一种线性空间是内积空间.

1. 内积空间的定义

定义 2.1.6(内积、实内积空间) 对于实数域$\mathbb{F}=\mathbb{R}$上的线性空间$(X,+,\alpha\cdot)$，若存在映射$(x,y):X\times X\to\mathbb{R}$，满足

(1) $\forall x\in X\Rightarrow(x,x)\geqslant 0$；$(x,x)=0\Leftrightarrow x=0$； (正定性 positive definiteness)

(2) $(x,y)=(y,x)$； (对称性 symmetry)

(3) $(\alpha x+\beta y,z)=\alpha(x,z)+\beta(y,z)$,
$(x,\alpha y+\beta z)=\alpha(x,y)+\beta(x,z)$. (线性 linear)

则称(x,y)为 X 上的**(实)内积(inner)**，并称$(X,+,\alpha\cdot)$为**实内积空间(real inner space)**，记为$(X,+,\alpha\cdot,(x,y))$. 也称其为**实欧几里得空间(real Euclidean space)**，并称映射(x,y)：$X\times X\to\mathbb{R}$ 为**双线性式(bilinear form)**.

定义 2.1.7(复内积、复内积空间) 对于复数域$\mathbb{F}=\mathbb{C}$上的线性空间$(X,+,\alpha\cdot)$，若存在映射$(x,y):X\times X\to\mathbb{C}$，满足

(1) $\forall x\in X\Rightarrow(x,x)\geqslant 0$；$(x,x)=0\Leftrightarrow x=0$； (正定性 positive definiteness)

(2) $(x,y)=\overline{(y,x)}$； (共轭对称 conjugate symmtry)

(3) $(\alpha x+\beta y,z)=\alpha(x,z)+\beta(y,z)$, (线性 linear)
$(x,\alpha y+\beta z)=\bar{\alpha}(x,y)+\bar{\beta}(x,z)$. (共轭线性 conjugate linear)

则称(x,y)为 X 上的**(复)内积**. $(X,+,\alpha\cdot,(x,y))$称为**复内积空间(complex inner space)**，或称其为**复欧几里得空间(complex Euclidean space)**.

注意，映射$(x,y):X\times X\to\mathbb{C}$ 不是双线性式.

例 2.1.7 实 n 维欧氏空间$(\mathbb{R}^n,+,\alpha\cdot)$，$\mathbb{R}^n=\{x=(x_1,\cdots,x_n):x_j\in\mathbb{R},1\leqslant j\leqslant n\}$，是最熟悉的内积空间，其元素间的内积为

$$(x,y)=\sum_{j=1}^{n}x_jy_j,\quad x,y\in\mathbb{R}^n.$$

例 2.1.8 $l^2=\left\{x=(x_1,x_2,\cdots,x_n,\cdots):\sum_{j=1}^{+\infty}x_j^2<+\infty\right\}$是内积空间，其元素间的内积为

$$(x,y)=\sum_{j=1}^{+\infty}x_jy_j,\quad x,y\in l^2.$$

例 2.1.9 $L^2([a,b])=\left\{f:\left(\int_a^b|f(x)|^2\mathrm{d}x\right)^{\frac{1}{2}}<+\infty\right\}$是内积空间，其元素间的内积为

$$(f,g)=\int_a^b f(x)g(x)\mathrm{d}x,\quad f,g\in L^2([a,b]).$$

2. 内积空间中元的长度、距离

在线性空间$(X,+,\alpha\cdot)$中，只有元之间的运算，却无法考虑元的“长度”. 然而，在内积空间中，内积$(x,y):X\times X\to\mathbb{R}(\mathbb{C})$却将“形”与“数”联系起来，赋予线性空间以更深层次的结构，这在以后的课程中会遇到很多例子. 现在以内积空间为例，先让读者认识集合的另一种结构 —— 拓扑结构的引进与意义.

定义 2.1.8(长度、距离) 在内积空间$(X,+,\alpha\cdot,(x,y))$中，对任意的$x\in X$，称$\|x\|=\sqrt{(x,x)}$为x的**长度(length)**；$\forall x,y\in X$，称$d(x,y)=\sqrt{(x-y,x-y)}$为x与y之间的**距离(distance)**.

有了距离的概念，就可以像高等数学中那样，引进“邻域”、“开集”、“闭集”、“序列的收敛性”、“极限”、“连续性”等一系列概念，向人们展示数学科学理论与应用更广阔的天地.

内积的下述性质很重要，会经常用到.

定理 2.1.6 内积(x,y)满足

(1) Cauchy 不等式 $|(x,y)|\leqslant\|x\|\cdot\|y\|,\quad x,y\in X$;

(2) 三角不等式 $\|x+y\|\leqslant\|x\|+\|y\|,\quad x,y\in X$;

(3) 距离的三点不等式 $d(x,y)\leqslant d(x,z)+d(z,y),\quad x,y,z\in X$;

(4) 平行四边形法则 $\|x+y\|^2+\|x-y\|^2=2\|x\|^2+2\|y\|^2,\quad x,y\in X$.

证明留作习题.

2.1.5 对偶空间

1. 线性空间上的映射

为了研究线性空间的结构，一个重要的概念是两个集合之间的“映射”. 在 1.1.3 节中，定义了两个集合之间的映射、满射、单射、逆映射、复合；进而，当集合自身有了运算结构后，如群、域等，在 1.2.2 中，又定义了两个群之间的同态、同构. 其实，在两个定义了运算结构

的集合上，都可以定义同态映射与同构映射.

同态映射，就是保持运算的映射，例如

群的同态 $f:(G_1,(\cdot)_{G_1})\to(G_2,(\cdot)_{G_2})$ 满足

$$f(x(\cdot)_{G_1}y)=f(x)(\cdot)_{G_2}f(y),\quad x,y\in G_1,\quad f(x),f(y)\in G_2;$$

域的同态 $f:(F_1,(+)_{F_1},(\cdot)_{F_1})\to(F_2,(+)_{F_2},(\cdot)_{F_2})$ 满足

$$f(x\,(+)_{F_1}y)=f(x)\,(+)_{F_2}f(y),\quad x,y\in F_1,\quad f(x),f(y)\in F_2,$$

$$f(x\,(\cdot)_{F_1}y)=f(x)\,(\cdot)_{F_2}f(y),\quad x,y\in F_1,\quad f(x),f(y)\in F_2;$$

线性空间的同态 $f:(X,(+)_X,(\alpha\cdot)_X)\to(Y,(+)_Y,(\alpha\cdot)_Y)$满足

$$f(x_1(+)_X x_2)=f(x_1)(+)_Y f(x_2),\quad x_1,x_2\in X,\ y_1=f(x_1),\ y_2=f(x_2)\in Y,$$

$$f((\alpha\cdot)_X x)=(\alpha\cdot)_Y f(x),\quad \alpha\in\mathbb{F},\ x\in X,\ y=f(x)\in Y.$$

事实上，线性空间之间的同态就是线性空间$(X,(+)_X,(\alpha\cdot)_X)$到$(Y,(+)_Y,(\alpha\cdot)_Y)$上的线性映射. 我们重新给出线性空间之间的线性映射的定义.

定义 2.1.9（线性映射） 数域$\mathbb{F}$上的线性空间$(X,(+)_X,(\alpha\cdot)_X)$到$(Y,(+)_Y,(\alpha\cdot)_Y)$的映射 $T:X\to Y$ 称为**线性映射(linear mapping)**，若 $T:X\to Y$ 满足

$$T(\alpha x_1+\beta x_2)=\alpha T(x_1)+\beta T(x_2),\quad \alpha,\beta\in\mathbb{F},\ x_1,x_2\in X.$$

映射，是函数概念的推广. 回顾在高等数学中的"函数"概念：函数 $f:\mathfrak{D}\to\mathfrak{R}$ 是从它的定义域 $\mathfrak{D}=[a,b]\subset\mathbb{R}$ 到值域 $\mathfrak{R}\subset\mathbb{R}$ 的一个**对应关系** f，按照这个 f，每个 $x\in\mathfrak{D}$ 仅有一个 $y\in\mathfrak{R}\subset\mathbb{R}$ 与之相对应. 上述同态与同构、线性映射等都是某种对应关系. 在自然科学领域中，考虑更广泛的集合之间、空间之间的自变量与因变量的"关系"时，便需要推广函数的概念. 通常，对于一个对应关系$f:\mathfrak{D}\to\mathfrak{R}$，

若 $\mathfrak{D}=[a,b]\subset\mathbb{R}$，$\mathfrak{R}\subset\mathbb{R}$，则称 f 为**函数(function)**；

若 $\mathfrak{D}\subset\mathbb{R}^n$，$\mathfrak{R}\subset\mathbb{R}$，则称 f 为**多元函数(several variable function)**；

若 $\mathfrak{D}\subset\mathbb{R}^n$，$\mathfrak{R}\subset\mathbb{R}^n$，则称 $\boldsymbol{f}$ 为**向量值函数(vector valued function)**；

若 $\mathfrak{D}\subset X$，$\mathfrak{R}\subset\mathbb{R}$，则称 f 为**泛函(functional)**；

若 $\mathfrak{D}\subset X$，$\mathfrak{R}\subset Y$，则称 f 为**映射(mapping)**，或**变换(transformation)**，或**算子(operator)**. 这里 X,Y 可以是群、域、线性空间，也可以是更一般的集合.

下面讨论一类新型的线性空间，其中的元是线性空间 X 上的线性泛函，称为 X 的对偶空间(或称共轭空间). 数域 $\mathbb{F}$ 均指实数域 $\mathbb{R}$ 或复数域 $\mathbb{C}$.

2. 对偶空间

1) 线性空间上的线性泛函、对偶空间

定义 2.1.10（X 上的线性泛函，对偶空间） 设$(X,+,\alpha\cdot)$为数域 $\mathbb{F}$($\mathbb{R}$ 或$\mathbb{C}$)上的线性空间，若映射 $f:X\to\mathbb{F}$ 满足 $f(\alpha x+\beta y)=\alpha f(x)+\beta f(y)$，$x,y\in X$，$\alpha,\beta\in\mathbb{F}$，则称 f 为 X 上的**线性泛函(linear functional)**.

记线性空间$(X,+,\alpha\cdot)$上线性泛函的全体为

$$X^* = \{f : f \text{ 是 } X \text{ 上的线性泛函}\}.$$

在X^*上定义加法运算$(+)_{X^*}$与数乘$(\alpha\cdot)_{X^*}$为

$$f,g \in X^* \Rightarrow (f\,(+)_{X^*}\,g)(x) = f(x)\,(+)_{\mathbb{F}}\,g(x), \quad x \in X;$$

$$f \in X^*,\ \alpha \in \mathbb{F} \Rightarrow ((\alpha\cdot)_{X^*}\,f)(x) = (\alpha\cdot)_{\mathbb{F}} f(x) = \alpha f(x), \quad x \in X,$$

则$(X^*,(+)_{X^*},(\alpha\cdot)_{X^*})$成为$\mathbb{F}$上的线性空间，并称$X^*$为$X$的**对偶空间(dual space)**，或**共轭空间(conjugate space)**.

例 2.1.10 取线性空间$(X,+,\alpha\cdot)$为三维(实)欧氏空间$\mathbb{R}^3$，我们证明$(\mathbb{R}^3)^*=\mathbb{R}^3$，这里等号理解成两个线性空间是同构的(亦即，元素之间一一对应、对应关系是保持两个空间运算的满射).

事实上，任取$y\in\mathbb{R}^3$，定义一个泛函$T_y:\mathbb{R}^3\to\mathbb{R}$，满足

$$T_y(x) = (x,y) = \sum_{j=1}^{3} x_j y_j, \quad x \in \mathbb{R}^3, \quad y \in \mathbb{R}^3,$$

其中$x=(x_1,x_2,x_3)\in\mathbb{R}^3$，$y=(y_1,y_2,y_3)\in\mathbb{R}^3$，$(x,y)$是$\mathbb{R}^3$中的内积. 不难验证，$\forall\, y\in\mathbb{R}^3$，有

$$T_y(\alpha x + \beta x') = \alpha T_y(x) + \beta T_y(x'), \quad \alpha,\beta \in \mathbb{R}, \quad x,x' \in \mathbb{R}^3.$$

于是有$\mathbb{R}^3\subset(\mathbb{R}^3)^*$.

反之，可证明$(\mathbb{R}^3)^*\subset\mathbb{R}^3$. 事实上，若$f:\mathbb{R}^3\to\mathbb{R}$是$\mathbb{R}^3$上的任一个线性泛函，即$f\in(\mathbb{R}^3)^*$，则$f(x+x')=f(x)+f(x')$与$f(\alpha x)=\alpha f(x)$确定了$f$为$x_1,x_2,x_3$的线性函数，因此必为$f(x)=f(x_1,x_2,x_3)=\sum_{k=1}^{3}c_k x_k$的形式，其中$c_k\in\mathbb{R}\ (k=1,2,3)$. 于是，任一线性泛函$f\in(\mathbb{R}^3)^*$必对应一组数$c_k\in\mathbb{R}\ (k=1,2,3)$，亦即对应于$\mathbb{R}^3$中的一个元，$f\in(\mathbb{R}^3)^* \xleftrightarrow{\text{对应}} (c_1,c_2,c_3)\in\mathbb{R}^3$. 从而$(\mathbb{R}^3)^*\subset\mathbb{R}^3$. 故$(\mathbb{R}^3)^*$与$\mathbb{R}^3$同构. 这里“$(\mathbb{R}^3)^*=\mathbb{R}^3$”中的等号是指$(\mathbb{R}^3)^*$与$\mathbb{R}^3$同构.

定义 2.1.11(线性泛函的零化子) 设$(X,+,\alpha\cdot)$为数域$\mathbb{F}$上的线性空间，$A\subset X$为X的非空子集，称对偶空间X^*中的子集$A^0=\{f\in X^*: f(A)=0\}$为**集合A的零化子(annihilator)**，其中$f(A)=\{f(x)\in\mathbb{F}: x\in A\}$是$A$的像集.

关于一个集合$A\subset X$的零化子，有如下性质：

(1) $A^0\subset X^*$是对偶空间X^*的线性子空间，无论集$A\subset X$是否是X的子空间. 当集$A\subset X$是X的子空间时，若$\dim X=n$有限，则$\dim A+\dim A^0=n$；

(2) 若集$A\subset B\subset X$，则$B^0\subset A^0$；

(3) 若$\dim X=n$为有限维，并视X^{**}与X等同，则对任一子集$A\subset X$，有$(A^0)^0=\operatorname{span}(A)$. 若$A$为$X$的子空间，则$(A^0)^0=A$；

(4) 若 $\dim X=n$ 有限，A,B 为子空间，则

$$(A\cap B)^0=A^0\cup B^0,\quad (A+B)^0=A^0\cap B^0;$$

(5) 若 $X=A\oplus B$ 为两个子空间 A 与 B 的直和，则

$$X^*=(A\oplus B)^*=A^0\oplus B^0.$$

2) 线性空间的二次对偶

定义 2.1.12（**X 的二次对偶空间**） 设$(X,+,\alpha\cdot)$为数域$\mathbb{F}$上的线性空间，X 的对偶空间为$(X^*,(+)_{X^*},(\alpha\cdot)_{X^*})$，则 X^* 也是$\mathbb{F}$上的线性空间. 定义空间 X^* 的对偶空间$(X^*)^*=X^{**}$，称为 X 的二次对偶. 三次对偶、四次对偶等可类似定义.

关于有限维线性空间的对偶空间有下面定理.

定理 2.1.7 设$(X,+,\alpha\cdot)$为数域$\mathbb{F}$上的 n 维线性空间，记 $\dim X=n$，则

(1) $\dim X=\dim X^*=\dim X^{**}=n$;

(2) X,X^*,X^{**} 彼此同构，并且对于 $x\in X$ 与 $x^{**}\in X^{**}$，有

$$x^{**}(w)=w(x),\quad \forall w\in X^*. \tag{2.1.1}$$

证 由假设，$\dim X=n$. 设 $\mathfrak{B}=\{e_1,e_2,\cdots,e_n\}\subset X$ 是 X 的一组基. 定义 X 上的n 个线性泛函 $e_k^*\in X^*(k=1,2,\cdots,n)$，使其满足

$$e_k^*(e_j)=\delta_{jk}=\begin{cases}1, & k=j,\\ 0, & k\neq j,\end{cases}\quad k,j=1,2,\cdots,n, \tag{2.1.2}$$

δ_{jk}称为**克罗内克(Kronecker)符号**.

易证，$\mathfrak{B}^*=\{e_1^*,e_2^*,\cdots,e_n^*\}\subset X^*$ 是 X^* 的一组基，称 $\mathfrak{B}^*$ 为 $\mathfrak{B}$的**对偶基(dual basis)**.

类似地，定义二次对偶基为 $\mathfrak{B}^{**}=\{e_1^{**},e_2^{**},\cdots,e_n^{**}\}\subset X^{**}$，使得

$$e_i^{**}(e_k^*)=\delta_{ki}=\begin{cases}1, & i=k,\\ 0, & i\neq k,\end{cases}\quad i,k=1,2,\cdots,n. \tag{2.1.3}$$

同样可证，$\mathfrak{B}^{**}=\{e_1^{**},e_2^{**},\cdots,e_n^{**}\}\subset X^{**}$ 是 X^{**} 的一组基，称 $\mathfrak{B}^{**}$ 为 $\mathfrak{B}$的二次对偶基.

于是，结论(1) $\dim X=\dim X^*=\dim X^{**}=n$ 得证.

为证(2)，定义映射

$$\tau_1: x=\sum_{j=1}^n\alpha_je_j\in X\to x^*=\sum_{j=1}^n\alpha_je_j^*\in X^*,\quad \alpha_j\in\mathbb{F},\quad j=1,2,\cdots,n,$$

显然，$\tau_1:X\to X^*$ 是同构映射，称为 X 到 X^* 的**自然同构(natural isomorphism)**. 将 $x=\sum_{j=1}^n\alpha_je_j\in X$ 记为 $[x]_{\mathfrak{B}}=\begin{bmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{bmatrix}$，将 $x^*=\sum_{j=1}^n\alpha_je_j^*\in X^*$ 记为 $[x^*]_{\mathfrak{B}^*}=\begin{bmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{bmatrix}$，则 X 与 X^* 同构，即

$$[x]_{\mathfrak{B}}=\begin{bmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{bmatrix}\xleftrightarrow{\tau_1}[x^*]_{\mathfrak{B}^*}=\begin{bmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{bmatrix}.$$

同理,定义 X^* 到 X^{**} 上的映射

$$\tau_2: x^*=\sum_{j=1}^{n}\alpha_j e_j^* \in X^* \to x^{**}=\sum_{j=1}^{n}\alpha_j e_j^{**}\in X^{**},\quad \alpha_j\in\mathbb{F}, j=1,2,\cdots,n,$$

τ_2 是 X^* 到 X^{**} 的自然同构,有

$$[x^*]_{\mathfrak{B}^*}=\begin{bmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{bmatrix}\xleftrightarrow{\tau_2}[x^{**}]_{\mathfrak{B}^{**}}=\begin{bmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{bmatrix}.$$

再来看 X 与 X^{**} 的关系. 令 $\tau=\tau_2\circ\tau_1: X\to X^{**}$,设 $x\in X$ 在基 $\mathfrak{B}$之下的表示为 $x=\sum_{j=1}^{n}\alpha_j e_j\in X$,下面来求 $\tau_2(\tau_1(x))$ 在基 $\mathfrak{B}^{**}$ 下的表示. 由

$$\tau(x)=\tau_2(\tau_1(x))=\tau_2(x^*)=x^{**},\tag{2.1.4}$$

任取 $w\in X^*$,设 w 在基 $\mathfrak{B}$ 之下的表示为

$$w=\sum_{i=1}^{n}\beta_i e_i^*\in X^*.\tag{2.1.5}$$

由(2.1.2)式知,$w(e_j)=\sum_{i=1}^{n}\beta_i e_i^*(e_j)=\beta_j, j=1,2,\cdots,n$. 将此式代入(2.1.5)式,得 $w=\sum_{i=1}^{n}w(e_i)e_i^*$. 这样,对于一个 $x^{**}=\sum_{j=1}^{n}\alpha_j e_j^{**}\in X^{**}$ 与任意 $w=\sum_{i=1}^{n}\beta_i e_i^*=\sum_{i=1}^{n}w(e_i)e_i^*\in X^*$,有

$$\begin{aligned}x^{**}(w)&=\left(\sum_{j=1}^{n}\alpha_j e_j^{**}\right)(w)=\sum_{j=1}^{n}\alpha_j e_j^{**}(w)=\sum_{j=1}^{n}\alpha_j e_j^{**}\left(\sum_{i=1}^{n}w(e_i)e_i^*\right)\\&=\sum_{j=1}^{n}\sum_{i=1}^{n}\alpha_j w(e_i)e_j^{**}(e_i^*)=\sum_{j=1}^{n}\alpha_j w(e_j)=w\left(\sum_{j=1}^{n}\alpha_j e_j\right)=w(x).\end{aligned}$$

这正是$[x^{**}]_{\mathfrak{B}^{**}}=\begin{bmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{bmatrix}\xleftrightarrow{\tau_2\circ\tau_1}[x]_{\mathfrak{B}}=\begin{bmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{bmatrix}$. 于是得到重要的对应关系

$$x^{**}=\sum_{j=1}^{n}\alpha_j e_j^{**}\in X^{**}\xleftrightarrow{\tau_2\circ\tau_1}x=\sum_{j=1}^{n}\alpha_j e_j\in X,$$

并成立等式(2.1.1). 归纳为表 2.1.1 中所示.

表 2.1.1

线性空间 X 基 $\mathfrak{B}=\{e_1,e_2,\cdots,e_n\}$	对偶空间 X^* 对偶基 $\mathfrak{B}=\{e_1^*,e_2^*,\cdots,e_n^*\}$	二次对偶空间 X^{**} 二次对偶基 $\mathfrak{B}^*=\{e_1^{**},e_2^{**},\cdots,e_n^{**}\}$
$e_k^*(e_j)=\delta_{jk}$		$e_i^{**}(e_k^*)=\delta_{ki}$
自然同构 $\tau_1:x=\sum_{j=1}^{n}\alpha_j e_j\in X$ $\to x^*=\sum_{j=1}^{n}\alpha_j e_j^*\in X^*$	自然同构 $\tau_2:x^*=\sum_{j=1}^{n}\alpha_j e_j^*\in X^*$ $\to x^{**}=\sum_{j=1}^{n}\alpha_j e_j^{**}\in X^{**}$	自然同构 $\tau=\tau_2\circ\tau_1:x=\sum_{j=1}^{n}\alpha_j e_j\in X$ $\to x^{**}=\sum_{j=1}^{n}\alpha_j e_j^{**}\in X^{**}$

2.1.6 线性空间的结构

线性空间$(X,+,\alpha\cdot)$的对偶空间与线性空间的结构有密切联系，特别是线性空间上的一种“双线性式”，就是 $X\times X$ 上的双线性泛函，它能决定线性空间$(X,+,\alpha\cdot)$的结构.

1. 线性空间上的双线性式、二次型

定义 2.1.13(**X 上的双线性式、二次型**) 设$(X,+,\alpha\cdot)$为数域$\mathbb{F}$上的线性空间，二元函数 $f:X\times X\to\mathbb{F}$ 称为 X 上的**双线性式**(**bilinear form**)，若 f 关于两个变量都是线性的，亦即

$$f(\alpha x+\beta y,z)=\alpha f(x,z)+\beta f(y,z),\quad x,y,z\in X,\quad \alpha,\beta\in\mathbb{F}$$

与

$$f(x,\alpha y+\beta z)=\alpha f(x,y)+\beta f(x,z),\quad x,y,z\in X,\quad \alpha,\beta\in\mathbb{F}.$$

将双线性式记为$(x,y):X\times X\to\mathbb{F}$. 称$(x,x):X\times X\to\mathbb{F}$ 为**二次型**(**quadratic form**).

定义 2.1.14(**对称与斜对称双线性式**) 设$(X,+,\alpha\cdot)$为数域$\mathbb{F}$上的线性空间，双线性式$(x,y):X\times X\to\mathbb{F}$ 称为**对称的**(**symmetric**)，若 $\forall x,y\in X\Rightarrow(x,y)=(y,x)$；称为**斜对称的**(**skew-symmetric**)(或称反对称的)，若 $\forall x,y\in X\Rightarrow(x,y)=-(y,x)$.

定义 2.1.15(**度量线性空间、非奇异度量线性空间**) 若在线性空间$(X,+,\alpha\cdot)$上定义了一个双线性式$(x,y):X\times X\to\mathbb{F}$，则称$(X,(x,y))\equiv(X,+,\alpha\cdot,(x,y))$为**度量线性空间**(**metric linear space**)；双线性式(x,y)称为度量线性空间$(X,(x,y))$上的**度量**.

对于度量线性空间$(X,(x,y))$，若 $\forall y\in X,(x,y)=0\Rightarrow x=0$，则称度量线性空间$(X,(x,y))$为**非奇异的**(**non-singular**).

2. 线性空间上的正交几何与辛几何

定义 2.1.16(**正交几何与辛几何**) 给定非奇异度量线性空间$(X,(x,y))$.

(1) 若$(x,y)=0$，则称 $x\in X$ 与 $y\in X$ **正交**(**orthogonal**)，也称直交，记为 $x\perp y$. 两个子

空间 $V,W\subset X$ 称为正交的，若 $\forall x\in V,y\in W$，都有 $x\perp y$，记为 $V\perp W$；

(2) 若 $(x,y)=(y,x)$，则称 $(X,(x,y))$ 为**对称度量线性空间**(**symmtric metric linear space**)，或称 X 是$\mathbb{F}$上的**正交几何**(**orthogonal geometry**)，也称**对称几何**；

(3) 若 $(x,y)=-(y,x)$，则称 $(X,(x,y))$ 为**斜对称度量线性空间**(**skew-symmtric metriclinear space**)，或称 X 是$\mathbb{F}$上的**辛几何**(**symplectic geometry**)，也称**斜对称几何**.

3. Riesz 表现定理

有限维非奇异度量线性空间上的线性泛函的表示定理，称为 **Riesz 表现定理**.

定理 2.1.8 设 $(X,(x,y))$ 是数域$\mathbb{F}$上的有限维非奇异度量线性空间，则对任一线性泛函 $f\in X^*$，存在惟一的元 $x\in X$，使得 $\forall v\in X$，泛函 f 对 $v\in X$ 的"作用"可表示为 $f(v)\equiv(v,x)$，式中的 (v,x) 是 X 的度量(双线性式).

证 取定一个 $x\in X$，定义映射 $\varphi_x:X\to\mathbb{F}$ 为 $\varphi_x(v)=(v,x)$. 显然，φ_x 是 X 上的线性泛函，亦即 $\varphi_x\in X^*$ 满足 $(\varphi_x,\alpha v_1+\beta v_2)=(\alpha v_1+\beta v_2,x)=\alpha(v_1,x)+\beta(v_2,x)$. 由此，再定义一个映射 $\tau:X\to X^*$，满足 $\tau(x)=\varphi_x$.

下面证明 $\tau:X\to X^*$ 是线性的. 对于 $x,y\in X$，$\alpha,\beta\in\mathbb{F}$，有

$$\tau(\alpha x+\beta y)=\alpha\tau(x)+\beta\tau(y). \tag{2.1.6}$$

事实上，对于任意的 $v\in X$，由

$$\begin{aligned}(\tau(\alpha x+\beta y),v)&=\varphi_{\alpha x+\beta y}(v)=(v,\alpha x+\beta y)=\alpha(v,x)+\beta(v,y)\\&=\alpha\varphi_x(v)+\beta\varphi_y(v)=(\alpha\varphi_x+\beta\varphi_y)(v)\\&=(\alpha\tau(x)+\beta\tau(y))(v),\end{aligned}$$

这就是 τ 的线性，即(2.1.6)式.

由于度量 (x,y) 是非奇异的，故集合 $\{x\in X:\varphi_x=0\}=\{x\in X:(v,x)=0,\forall v\in X\}$ 是线性空间 X 中只含零元的集合 $\{x\in X:\varphi_x=0\}=\{0\}$，因此，映射 $\tau:X\to X^*$ 是一对一的满映射，所以它是 X 到 X^* 上的同构. 于是，当 X 是有限维($\dim X=n$)时，其上的线性泛函 $\varphi_x\in X^*$ 一定是 $\varphi_x(v)=(v,x)$. 定理得证.

Riesz 表现定理告诉我们，有限维非奇异度量线性空间 $(X,+,\alpha\cdot,(x,y))$ 上的线性泛函只有一种，就是定义在此空间上的双线性式.

4. 线性空间上的正交几何与辛几何的结构

1) 线性空间的变换矩阵

研究有限维线性空间 X 的两组基之间的关系至关重要. 下面从回顾$\mathbb{R}^3$ 中的情形开始.

(1) $\mathbb{R}^3$ 中两个坐标系的变换公式

回忆在高等数学中，讨论了$\mathbb{R}^3$ 中的直角坐标系 $Oxyz$ 与仍以原点 O 为新坐标原点的新坐标系 $Ox_1y_1z_1$ 之间的关系. 记 $\mathfrak{B}=\{Oxyz:i,j,k\}$ 与 $\mathfrak{C}=\{Ox_1y_1z_1:i_1,j_1,k_1\}$，其中 $\{i,j,k\}$，

$\{i_1,j_1,k_1\}$分别为两个坐标系中坐标轴上的正交单位向量,也就是$\mathbb{R}^3$中的两组基,其间的夹角由表2.1.2给出. 基的变换公式写成线性变换(方程组)的形式为

$$\begin{cases} i_1 = i\cos\alpha_1 + j\cos\alpha_2 + k\cos\alpha_3, \\ j_1 = i\cos\beta_1 + j\cos\beta_2 + k\cos\beta_3, \\ k_1 = i\cos\gamma_1 + j\cos\gamma_2 + k\cos\gamma_3, \end{cases} \tag{2.1.7}$$

这里的变换方阵 $C=\begin{bmatrix} \cos\alpha_1 & \cos\alpha_2 & \cos\alpha_3 \\ \cos\beta_1 & \cos\beta_2 & \cos\beta_3 \\ \cos\gamma_1 & \cos\gamma_2 & \cos\gamma_3 \end{bmatrix}$是正交方阵,$CC^{\mathrm{T}}=I$,$C^{\mathrm{T}}$是$C$的转置方阵.

表 2.1.2

	i_1	j_1	k_1
i	α_1	β_1	γ_1
j	α_2	β_2	γ_2
k	α_3	β_3	γ_3

令 $\delta=\begin{bmatrix} i \\ j \\ k \end{bmatrix}$,$\delta_1=\begin{bmatrix} i_1 \\ j_1 \\ k_1 \end{bmatrix}$,则基的变换公式(2.1.7)可写为$\delta_1=C\delta$. 记

$$\mathfrak{B}\equiv\{i,j,k\}\leftrightarrow M_{\mathfrak{B}}\equiv\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}\in\mathfrak{M}_1,\quad \mathfrak{C}\equiv\{i_1,j_1,k_1\}\leftrightarrow M_{\mathfrak{C}}\equiv\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}\in\mathfrak{M}_1.$$

于是,$\mathfrak{B}=\{i,j,k\}\xrightarrow{C}\mathfrak{C}=\{i_1,j_1,k_1\}$.(注意到,基$\mathfrak{B}$本来对应于一个$3\times3$的方阵,这里隐去每个三维向量$i,j,k$的分量,将它们直接视为$\mathbb{R}^3$中的三个元(行向量)$b_1,b_2,b_3$,因此$\mathfrak{B}\to M_{\mathfrak{B}}\in\mathfrak{M}_1$,这在研究高维情形时将是很方便的.)

记$C\equiv M_{\mathfrak{B}}$,$M_{\mathfrak{B}}:M_{\mathfrak{B}}\to M_{\mathfrak{C}}$称为**变换方阵**,则$\delta_1=C\delta$表示为**基的变换公式**

$$M_{\mathfrak{C}}=M_{\mathfrak{B}}M_{\mathfrak{B}} \tag{2.1.8}$$

为确定变换方阵$C\equiv M_{\mathfrak{B}}\in\mathfrak{M}_3$,记$C\equiv[C_1\ C_2\ C_3]$,将变换写为

$$[C(M_{\mathfrak{B}})]_{\mathfrak{C}}=CM_{\mathfrak{B}}$$

取$\mathfrak{B}=\mathfrak{N}=\{e_1,e_2,e_3\}$中的元为$e_1=\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$,$e_2=\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$,$e_3=\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$,则$[C(e_j)]_{\mathfrak{C}}=Ce_j=C_j\ (j=1,2,3)$,因此得到变换方阵为$M_{\mathfrak{B}}\equiv[C(e_j)]_{\mathfrak{C}}$.

(2) n维线性空间X与m维线性空间Y的坐标系的变换公式

上述(1)的结果可推广为n维线性空间X与m维线性空间Y的坐标系的变换.

设 X,Y 的维数分别为 $\dim X=n$ 与 $\dim Y=m$. X 的基 $\mathfrak{B}$对应于 $\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \equiv M_{\mathfrak{B}}$, Y 的基 $\mathfrak{C}$对应于 $\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{bmatrix} \equiv M_{\mathfrak{C}}$; 于是有对应关系: $\mathfrak{B}\leftrightarrow M_{\mathfrak{B}}\in \mathfrak{M}_{n\times 1}$ 与 $\mathfrak{C}\leftrightarrow M_{\mathfrak{C}}\in \mathfrak{M}_{m\times 1}$. 若线性变换 $\tau_{\mathfrak{BC}}:\mathfrak{B}\to\mathfrak{C}$将 X 的基 $\mathfrak{B}$变换到 Y 的基 $\mathfrak{C}$,则有

$$\tau_{\mathfrak{BC}}\leftrightarrow M_{\mathfrak{BC}} \in \mathfrak{M}_{m\times n}; \tag{2.1.9}$$

基的变换公式为

$$M_{\mathfrak{C}} = M_{\mathfrak{BC}}M_{\mathfrak{B}} \tag{2.1.10}$$

如(1)同样的推理可得到,**变换矩阵**为 $M_{\mathfrak{BC}}=[[\tau_{\mathfrak{BC}}(e_1)]_{\mathfrak{C}}\cdots[\tau_{\mathfrak{BC}}(e_n)]_{\mathfrak{C}}]$.

(3) $\mathbb{R}^3$ 中两个坐标系变换之下点的坐标变换公式

在$\mathbb{R}^3$ 中,设两个坐标系的变换为 $\tau_{\mathfrak{BC}}:\mathfrak{B}\to\mathfrak{C}$,点 $P\in\mathbb{R}^3$ 的坐标变换公式推导如下.

由 $\mathfrak{B}=\{i,j,k\}\leftrightarrow\delta\equiv\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}\in\mathfrak{M}_{3\times 1}$,点 P 在 $\mathfrak{B}$中的坐标表示为 $(x_1,x_2,x_3)\equiv[x]_{\mathfrak{B}}$; 又 $\mathfrak{C}=\{i_1,j_1,k_1\}\leftrightarrow\delta_1\equiv\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}\in\mathfrak{M}_{3\times 1}$,点 P 在 $\mathfrak{C}$中的坐标表示为 $(x_1',x_2',x_3')\equiv[x]_{\mathfrak{C}}$; 再由 $\tau_{\mathfrak{BC}}:\mathfrak{B}\to\mathfrak{C}$为正交变换,故表示矩阵为 $M_{\mathfrak{BC}}:\delta\to\delta_1\,(M_{\mathfrak{C}}=M_{\mathfrak{BC}}M_{\mathfrak{B}})$,且 $M_{\mathfrak{BC}}(M_{\mathfrak{BC}})^{\mathrm{T}}=I$, $(M_{\mathfrak{BC}})^{\mathrm{T}}=(M_{\mathfrak{BC}})^{-1}$ 为正交方阵. 注意到正交变换 $\tau_{\mathfrak{BC}}:\mathfrak{B}\to\mathfrak{C}$不改变$\overrightarrow{OP}$的长度,也不会改变任意两个元的内积,我们取原坐标系下的两个元 $[x]_{\mathfrak{B}}\equiv\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\in\mathfrak{M}_{3\times 1}$与 $\delta\equiv\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}\in\mathfrak{M}_{3\times 1}$,它们在新坐标系下的表示分别为

$$[x]_{\mathfrak{C}}\equiv\begin{bmatrix} x_1' \\ x_2' \\ x_3' \end{bmatrix}\in\mathfrak{M}_{3\times 1} \quad 与 \quad \delta_1\equiv\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}\in\mathfrak{M}_{3\times 1}.$$

作内积,有$[x]_{\mathfrak{C}}\cdot\delta_1=[x]_{\mathfrak{B}}\cdot\delta$,并推导如下:

$$[x]_{\mathfrak{C}}\cdot\delta_1 = [x]_{\mathfrak{B}}\cdot\delta \qquad (由\ \delta=(M_{\mathfrak{BC}})^{-1}\delta_1=(M_{\mathfrak{BC}})^{\mathrm{T}}\delta_1)$$

$$\Rightarrow\quad [x]_{\mathfrak{C}}\cdot\delta_1 = [x]_{\mathfrak{B}}\cdot\delta = [x]_{\mathfrak{B}}\cdot\{(M_{\mathfrak{BC}})^{\mathrm{T}}\delta_1\} \qquad (将此等式右乘(\delta_1)^{\mathrm{T}})$$

$$\Rightarrow\quad \{[x]_{\mathfrak{C}}\cdot\delta_1\}(\delta_1)^{\mathrm{T}} = \{[x]_{\mathfrak{B}}\cdot(M_{\mathfrak{BC}})^{\mathrm{T}}\delta_1\}(\delta_1)^{\mathrm{T}} \quad (直接验证)$$

$$\Rightarrow \quad [x]_{\mathfrak{C}} = M_{\mathfrak{B}}[x]_{\mathfrak{B}}$$

于是，$[x]_{\mathfrak{C}}=M_{\mathfrak{B}}[x]_{\mathfrak{B}}$就是$\mathbb{R}^3$ **中在两个坐标系变换下的一个点的坐标变换公式.**

由此可见，变换方阵 $M_{\mathfrak{B}}$不仅将$\mathfrak{B}$的基 $M_{\mathfrak{B}}$变为$\mathfrak{C}$的基 $M_{\mathfrak{C}}$($M_{\mathfrak{C}}=M_{\mathfrak{B}}M_{\mathfrak{B}}$)，也将一个点的坐标$[x]_{\mathfrak{B}}$变为$[x]_{\mathfrak{C}}$($[x]_{\mathfrak{C}}=M_{\mathfrak{B}}[x]_{\mathfrak{B}}$).

(4) n 维线性空间 X 与 m 维线性空间 Y 的坐标系变换下点的坐标变换公式

上述(3)的结果可推广为 n 维线性空间 X 与 m 维线性空间 Y 的坐标系变换下点的坐标变换公式. 在(2)的假设下，再设任一个元 $x\in X$ 在基 $\mathfrak{B}\subset X$ 的坐标表示为$[x]_{\mathfrak{B}}\equiv\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}\in \mathfrak{M}_{n\times 1}$，任一个元 $y\in Y$ 在基 $\mathfrak{C}\subset Y$ 的坐标表示为$[y]_{\mathfrak{C}}\equiv\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_m \end{bmatrix}\in \mathfrak{M}_{m\times 1}$.

于是，由变换 $\tau_{\mathfrak{B}}:\mathfrak{B}\to\mathfrak{C}$ 也有 $\tau_{\mathfrak{B}}:X\to Y$，使得 $\tau_{\mathfrak{B}}(x)=y$，并且

$$\tau_{\mathfrak{B}}([x]_{\mathfrak{B}}) = [y]_{\mathfrak{C}}, \tag{2.1.11}$$

或用变换 $\tau_{\mathfrak{B}}$的变换矩阵 $M_{\mathfrak{B}}$表示为

$$[y]_{\mathfrak{C}} = M_{\mathfrak{B}}[x]_{\mathfrak{B}} \tag{2.1.12}$$

2) 有限维非奇异度量线性空间

关于非奇异 n 维度量线性空间，我们还有如下定理.

定理 2.1.9 设$(X,+,\alpha\cdot,(x,y))$是数域$\mathbb{F}$上的有限维($\dim X=n$)非奇异度量线性空间，设 $\mathfrak{B}=\{b_1,b_2,\cdots,b_n\}$是 X 的基，则方阵$[\mathfrak{B}]\equiv[b_{jk}]_{n\times n}$在基

$$\mathfrak{B}=\{b_1,b_2,\cdots,b_n\}=\{[b_{11}\ \cdots\ b_{1n}]^{\mathrm{T}},\cdots,[b_{n1}\ \cdots\ b_{nn}]^{\mathrm{T}}\}$$

之下定义为$[\mathfrak{B}]=[(b_j,b_k)]_{n\times n}$，其中$[\because]^{\mathrm{T}}$为方阵$[\because]$的转置，称为由**双线性式**$(x,y)$**决定的对应方阵**，简称对应方阵. 则

(1) 双线性式(x,y)是对称的，当且仅当对应方阵$[\mathfrak{B}]=[(b_j,b_k)]_{n\times n}$是对称的，亦即 $b_{jk}=b_{kj}(1\leqslant j,k\leqslant n)$；

(2) 双线性式(x,y)是斜对称的，当且仅当对应方阵$[\mathfrak{B}]=[(b_j,b_k)]_{n\times n}$是斜对称的，亦即 $(b_j,b_j)=0,(b_j,b_k)=-(b_k,b_j)(1\leqslant j\neq k\leqslant n)$.

证 对于 $x\in X,y\in X$，在基 $\mathfrak{B}=\{b_1,b_2,\cdots,b_n\}$中的表示分别为 $x=\sum_{j=1}^{n}x_jb_j$，$y=$

$\sum_{j=1}^{n} y_j b_j$，并记为$[x]_{\mathfrak{B}}=\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$，$[y]_{\mathfrak{B}}=\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$，则双线性式在基 $\mathfrak{B}$之下的表示为

$$(x,y)=\left(\sum_{j=1}^{n} x_j b_j,\sum_{k=1}^{n} y_k b_k\right)=\sum_{j=1}^{n} x_j\left(b_j,\sum_{k=1}^{n} y_k b_k\right)=\sum_{j=1}^{n}\sum_{k=1}^{n} x_j y_k (b_j,b_k).$$

因此，对任意 $x,y\in X$，若$(x,y)=(y,x)$，则必蕴含$(b_j,b_k)=(b_k,b_j)(1\leqslant j,k\leqslant n)$，故$[\mathfrak{B}]=[(b_j,b_k)]_{n\times n}$是对称方阵；反之亦然.

当$(x,y)=-(y,x)$时，则有$(b_j,b_k)=-(b_k,b_j)(1\leqslant j,k\leqslant n)$，从而$(b_j,b_j)=0$，与$(b_j,b_k)=-(b_k,b_j)(1\leqslant j\neq k\leqslant n)$，故必蕴含$[\mathfrak{B}]$是斜对称方阵；反之亦然.

3）正交几何的分解定理

定理 2.1.10 设$(X,+,\alpha\cdot,(x,y))$是数域$\mathbb{F}$上的 $\dim X=n$ 维、非奇异、对称度量线性空间，则必存在正交基 $\mathfrak{B}=\{b_1,b_2,\cdots,b_n\}$，$(b_j,b_k)=\begin{cases}(b_j,b_j)\neq 0, & k=j,\\ 0, & k\neq j\end{cases}$ $(j,k=1,2,\cdots,n)$，使得 X 可分解为正交直和 $X=S_1\underline{\oplus}\cdots\underline{\oplus}S_n$，其中：

① 符号$\underline{\oplus}$表示**正交直和**，亦即 $S_j\perp S_k$，$S_j\oplus S_k$，$j\neq k$，且 $\forall x\in X$ 可表示为 $x=\sum_{j=1}^{n} s_j$，$s_j\in S_j$，$j=1,2,\cdots,n$；

② 由 b_j，$(b_j,b_j)=a_j$，$j=1,2,\cdots,n$ 生成的一维子空间S_j $(j=1,2,\cdots,n)$是 $S_j=\{x\in X: x=\alpha b_j, b_j\in\mathfrak{B}\}$；

③ 空间 X 的度量（双线性式）(x,y)关于正交基 $\mathfrak{B}=\{b_1,b_2,\cdots,b_n\}$的对应方阵$[\mathfrak{B}]=\begin{bmatrix}(b_1,b_1) & \cdots & (b_1,b_n)\\ \vdots & \ddots & \vdots \\ (b_n,b_1) & \cdots & (b_n,b_n)\end{bmatrix}$为**非奇异正交方阵**，$|[\mathfrak{B}]|\neq 0$，且 $([\mathfrak{B}])^{\mathrm{T}}[\mathfrak{B}]=[(b_j,b_j)]_{n\times n}$. 当数域$\mathbb{F}=\mathbb{C}$时（它是一个代数封闭域，即所有以$\mathbb{F}$中的元为系数的 n 次多项式都可分解为一次式的乘积），则必存在**标准正交基** $\mathfrak{N}$使得对应方阵$[\mathfrak{B}]$必可取为

$$[\mathfrak{B}]=\begin{bmatrix}1 & & \\ & \ddots & \\ & & 1\end{bmatrix};\tag{2.1.13}$$

当数域$\mathbb{F}=\mathbb{R}$时（它不是一个代数封闭域），则必存在**正交基** $\mathfrak{B}=\{b_1=(b_{11},\cdots,b_{1n}),\cdots,b_n=(b_{n1},\cdots,b_{nn})\}$满足 $(b_j,b_k)=\begin{cases}(b_j,b_j)\neq 0, & k=j,\\ 0, & k\neq j,\end{cases}$ 且 $(b_j,b_j)=\begin{cases}1, & 1\leqslant j\leqslant j_0,\\ -1, & j_0<j\leqslant n,\end{cases}$ 使对应方阵为

$$[\mathfrak{B}] = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & -1 & & \\ & & & & \ddots & \\ & & & & & -1 \end{bmatrix}; \qquad (2.1.14)$$

④ 若 $\boldsymbol{P}$ 是一个 n 阶非奇异对称方阵,则必存在 n 阶非奇异方阵 Q,使得

$$P = Q^{\mathrm{T}}[\mathfrak{B}]Q = Q^{\mathrm{T}} \begin{bmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{bmatrix} Q,$$

其中 $(b_j, b_j) = a_j$ 由 ② 确定;在 $\mathbb{F} = \mathbb{C}$ 时,Q 由(2.1.13)式给出;在 $\mathbb{F} = \mathbb{R}$ 时,Q 由(2.1.14)式给出.

4)斜几何的分解定理

定理 2.1.11 设 $(X, +, \alpha \cdot, (x, y))$ 是数域 $\mathbb{F}$ 上的 $\dim X = n$ 维、非奇异、斜对称度量线性空间,则必存在正交基 $\mathfrak{B} = \{b_1, b_2, \cdots, b_n\}$,使得 X 可分解为正交直和 $X = H_1 \underline{\oplus} \cdots \underline{\oplus} H_k$,其中:① 符号 $\underline{\oplus}$ 表示**正交直和**,亦即 $H_j \perp H_l$, $H_j \oplus H_l$, $j \neq l$,且 $\forall x \in X$ 可表示为 $x = \sum_{j=1}^{k} h_j$, $h_j \in H_j$, $j = 1, 2, \cdots, k$;

② H_j 是二维斜对称度量子空间,其对应方阵为 $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$;因此,非奇异、斜对称、度量线性空间必定是偶数维的;

③ 可选取一组正交基 $\mathfrak{B} = \{b_1, b_2, \cdots, b_n\}$,使得度量 (x, y) 的对应方阵为

$$[\mathfrak{B}] = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ -1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & -1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 0 & \cdots & -1 & 0 \end{bmatrix};$$

④ 若 P 是一个 n 阶非奇异斜对称方阵,则必存在 n 阶非奇异方阵 Q,使得

$$P=Q^{\mathrm{T}}[\mathfrak{B}]Q=Q^{\mathrm{T}}\begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 & 0 \\ -1 & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 & 0 \\ 0 & 0 & -1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & 0 & \cdots & -1 & 0 \end{bmatrix}Q,$$

因此,非奇异斜对称方阵也一定是偶数阶的.

5）惯性定理（sylvester's law of inertia）

定理 2.1.12 设$(X,+,\alpha\cdot,(x,y))$是数域$\mathbb{F}$上的$\dim X=n$维非奇异度量线性空间.

① 若(x,y)是对称的，则存在X的一组**正交基** $\mathfrak{B}=\{b_1,b_2,\cdots,b_n\}$，使得对于任意$x,y\in X$，有

$$(x,y)=a_1\xi_1\eta_1+a_2\xi_2\eta_2+\cdots+a_n\xi_n\eta_n, \tag{2.1.15}$$

若(x,y)是斜对称的，则存在X的一组**正交基** $\mathfrak{B}=\{b_1,b_2,\cdots,b_n\}$，使得对任意$x,y\in X$，有

$$(x,y)=\xi_1\eta_2-\xi_2\eta_1+\cdots+\xi_{n-1}\eta_n-\xi_n\eta_{n-1}, \tag{2.1.16}$$

(2.1.15)式、(2.1.16)式中，$[x]_{\mathfrak{B}}=\begin{bmatrix}\xi_1\\ \xi_2\\ \vdots\\ \xi_n\end{bmatrix}$，$[y]_{\mathfrak{B}}=\begin{bmatrix}\eta_1\\ \eta_2\\ \vdots\\ \eta_n\end{bmatrix}$，$(b_j,b_j)=a_j$，$j=1,2,\cdots,n$；

② 若数域$\mathbb{F}=\mathbb{C}$，则存在一组**标准正交基** $\mathfrak{N}=\{b_1=(b_{11},\cdots,b_{1n}),\cdots,b_n=(b_{n1},\cdots,b_{nn})\}$，$(b_j,b_k)=\begin{cases}1, & k=j,\\ 0, & k\neq j,\end{cases}$使得对于任意$x,y\in X$，有

$$(x,y)=\xi_1\eta_1+\xi_2\eta_2+\cdots+\xi_n\eta_n, \tag{2.1.17}$$

若数域$\mathbb{F}=\mathbb{R}$，则存在一组**正交基** $\mathfrak{N}=\{b_1=(b_{11},\cdots,b_{1n}),\cdots,b_n=(b_{n1},\cdots,b_{nn})\}$，$(b_j,b_k)=\begin{cases}(b_j,b_j)\neq 0, & k=j,\\ 0, & k\neq j,\end{cases}$且$(b_j,b_j)=\begin{cases}1, & 1\leqslant j\leqslant j_0,\\ -1, & j_0<j\leqslant n,\end{cases}$使得对于任意$x,y\in X$，有

$$(x,y)=\xi_1\eta_1+\cdots+\xi_k\eta_k-\xi_{k+1}\eta_{k+1}-\cdots-\xi_n\eta_n; \tag{2.1.18}$$

③ 特别地，对于二次型(x,x)，若数域$\mathbb{F}=\mathbb{C}$，则$(x,x)=\xi_1^2+\xi_2^2+\cdots+\xi_n^2$；若数域$\mathbb{F}=\mathbb{R}$，则$(x,x)=\xi_1^2+\cdots+\xi_k^2-\xi_{k+1}^2-\cdots-\xi_n^2$.

2.2 线性变换

2.2.1 线性算子空间

1. 线性空间上的线性算子

本节考虑同一个数域$\mathbb{F}$上的两个线性空间X,Y之间的线性变换.

定义 2.2.1（线性算子、线性算子空间） 设$T:X\to Y$是数域$\mathbb{F}$上的线性空间$(X,(+)_X,(\alpha\cdot)_X)$到线性空间$(Y,(+)_Y,(\alpha\cdot)_Y)$之间的映射，若$T$满足

$$T(\alpha x_1+\beta x_2)=\alpha T(x_1)+\beta T(x_2)=\alpha y_1+\beta y_2,\quad x_1,x_2\in X,\ y_1,y_2\in Y,\ \alpha,\beta\in\mathbb{F},$$

则称T为X到Y的**线性算子（linear operator）**（或称**线性映射（linear mapping）**，或**线性变换（linear transformation）**）.

线性算子空间　记X到Y的线性算子的全体为

$$L(X,Y)=\{T:X\to Y\text{ 是 }X\text{ 到 }Y\text{ 的线性算子}\},$$

记其中的元为 T,S,R，等. 在 $L(X,Y)$上定义加法运算与数乘运算：

加法 $+\equiv(+)_{L(X,Y)}\quad T,S\in L(X,Y)\Rightarrow(T+S)(x)=T(x)+S(x),\forall x\in X$；

数乘 $\alpha\cdot\equiv(\alpha\cdot)_{L(X,Y)}\quad T\in L(X,Y),\quad \alpha\in\mathbb{F}\Rightarrow(\alpha\cdot T)(x)=\alpha T(x),\forall x\in X,\quad \alpha\in\mathbb{F}$；

则 $L(X,Y)$成为数域$\mathbb{F}$ 上的线性空间，称为**线性算子空间**(**linear operator space**).

特别地，当 Y 是数域$\mathbb{F}$($\mathbb{R}$ 或$\mathbb{C}$)时，$L(X,\mathbb{F})$就是**线性泛函空间**(**linear functional space**)，亦即 X 的**对偶空间** $X^*=L(X,\mathbb{F})$(定义 2.1.10)，也称 X^* 为 X 的**共轭空间**.

我们有线性算子空间 $L(X,Y)$的同构定理.

定理 2.2.1 设 X,Y 分别为数域$\mathbb{F}$ 上的 $\dim X=n$ 维与 $\dim Y=m$ 维线性空间，则

$$L(X,Y)\xleftrightarrow{\text{同构}}\mathfrak{M}_{mn}\equiv\mathfrak{M}_{mn}(\mathbb{F}),$$

同构映射由对应关系 $T\in L(X,Y)\leftrightarrow A=[T]_{\mathfrak{B}}=[a_{ij}]_{1\leqslant i\leqslant m,1\leqslant j\leqslant n}$给出.

证 设 X 的一组基为 $\mathfrak{B}=\{b_1,b_2,\cdots,b_n\}$，每个 $x\in X$ 关于 $\mathfrak{B}$的坐标记为$[x]_{\mathfrak{B}}=\begin{bmatrix}\alpha_1\\\alpha_2\\\vdots\\\alpha_n\end{bmatrix}$；设 Y 的一组基为 $\mathfrak{C}=\{c_1,c_2,\cdots,c_m\}$，每个 $y\in Y$ 关于 $\mathfrak{C}$的坐标记为$[y]_{\mathfrak{C}}=\begin{bmatrix}\beta_1\\\beta_2\\\vdots\\\beta_m\end{bmatrix}$. 对于任一个线性算子 $T\in L(X,Y)$，设 T 将基 $\mathfrak{B}$映到基 $\mathfrak{C}$，则 $T:\mathfrak{B}\to\mathfrak{C}$，且 $T(x)=y\in Y$. 下面来确定 T.

取$[e_1]_{\mathfrak{B}}=\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix},\cdots,[e_n]_{\mathfrak{B}}=\begin{bmatrix}0\\0\\\vdots\\1\end{bmatrix}$，由$[T(x)]_{\mathfrak{C}}=[y]_{\mathfrak{C}}=\begin{bmatrix}\beta_1\\\beta_2\\\vdots\\\beta_m\end{bmatrix}$，则$[T(x)]_{\mathfrak{C}}=A\,[x]_{\mathfrak{B}}$ 为确定矩阵 A，令 $A=[A_1\ A_2\ \cdots\ A_n]$，其中 $A_j=\begin{bmatrix}a_{1j}\\a_{2j}\\\vdots\\a_{mj}\end{bmatrix}$，$j=1,2,\cdots,n$. 取标准正交基

$$\mathfrak{N}=\{b_1=(1,0,\cdots,0),\cdots,b_n=(0,\cdots,0,1)\},$$

亦即$[b_1]_{\mathfrak{B}}=[e_1]_{\mathfrak{B}},\cdots,[b_n]_{\mathfrak{B}}=[e_n]_{\mathfrak{B}}$，代入$[T(x)]_{\mathfrak{C}}=A\,[x]_{\mathfrak{B}}$，得到

$$[T(b_1)]_{\mathfrak{C}}=\begin{bmatrix}(T(b_1))_1\\ \vdots\\ (T(b_1))_m\end{bmatrix}=A[b_1]_{\mathfrak{B}}=[A_1\ A_2\ \cdots\ A_n][b_1]_{\mathfrak{B}}=[A_1\ A_2\ \cdots\ A_n]\begin{bmatrix}1\\0\\ \vdots\\0\end{bmatrix}=A_1.$$

对$[T(b_j)]_{\mathfrak{C}}(j=2,\cdots,n)$作同样处理,得到$[T(b_j)]_{\mathfrak{C}}=\begin{bmatrix}(T(b_j))_1\\ \vdots\\ (T(b_j))_m\end{bmatrix}_{\mathfrak{B}}=A[b_j]_{\mathfrak{B}}=A_j$. 于是,有

$A_j=[T(b_j)]_{\mathfrak{C}}(j=1,2,\cdots,n)$,故

$$A\equiv[T]_{\mathfrak{B}}=[A_1\ A_2\ \cdots\ A_n]=[[T(b_1)]_{\mathfrak{C}}\cdots[T(b_n)]_{\mathfrak{C}}]_{m\times n}.$$

因此,$T\in L(X,Y)\leftrightarrow A=[T]_{\mathfrak{B}}\in\mathfrak{M}_{m\times n}$.

于是,T 对应于一个矩阵,即 $T\in L(X,Y)\leftrightarrow[T]_{\mathfrak{B}}\in\mathfrak{M}_{m\times n}$. 并且对于任一个 $x\in X$,设其关于 $\mathfrak{B}$的坐标为 $[x]_{\mathfrak{B}}=\begin{bmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \alpha_n\end{bmatrix}$,而像点 $T(x)\in Y$ 关于 $\mathfrak{C}$的坐标为 $[T(x)]_{\mathfrak{C}}=\begin{bmatrix}\beta_1\\ \beta_2\\ \vdots\\ \beta_m\end{bmatrix}$; 则成立公式

$$[T(x)]_{\mathfrak{C}}=[T]_{\mathfrak{B}}[x]_{\mathfrak{B}},$$

其中 $A=[T]_{\mathfrak{B}}:\mathfrak{B}\to\mathfrak{C}$,从而 $T\in L(X,Y)$ 决定一个由$\mathbb{F}^n$ 与$\mathbb{F}^m$之间的线性变换

$$A=[T]_{\mathfrak{B}}\in\mathfrak{M}_{m\times n},$$

并且 $T\in L(X,Y)\to A=[T]_{\mathfrak{B}}\in L(\mathbb{F}^n,\mathbb{F}^m)\leftrightarrow\mathfrak{M}_{m\times n}$.

现在定义一个映射 $\varphi:L(X,Y)\to\mathfrak{M}_{m\times n}$,来证明 φ 是 $L(X,Y)$到 $\mathfrak{M}_{m\times n}$的同构,即

$$L(X,Y)\overset{\varphi}{\longleftrightarrow}\mathfrak{M}_{m\times n}\equiv\mathfrak{M}_{m\times n}(\mathbb{F}).$$

令 $\varphi(T)=A=[T(b_j)]_{\mathfrak{C}}$, $\forall\,T\in L(X,Y)$,它是线性映射,且是 $L(X,Y)$到 $\mathfrak{M}_{m\times n}$上的同构映射.

线性——事实上,$\forall\,\alpha,\beta\in\mathbb{F}$,$\forall\,T,S\in L(X,Y)$,根据定义有

$$\varphi(\alpha T+\beta S)=[\alpha T(b_j)+\beta S(b_j)]_{\mathfrak{C}}=\alpha[T(b_j)]_{\mathfrak{C}}+\beta[S(b_j)]_{\mathfrak{C}}=\alpha\varphi(T)+\beta\varphi(S).$$

满射——对于任一个矩阵 $B\in\mathfrak{M}_{m\times n}$,都可写成 $B=[B_1\ B_2\ \cdots\ B_n]$,$B_j=\begin{bmatrix}\beta_{1j}\\ \beta_{2j}\\ \vdots\\ \beta_{mj}\end{bmatrix}$,$j=1,2,\cdots,n$. 对于确定的基 $\mathfrak{B}=\{b_1,b_2,\cdots,b_n\}$,只需选取一个映射 $T\in L(X,Y)$,使得$[T(b_j)]_{\mathfrak{C}}=B_j$ $(j=1,2,\cdots,n)$,这是一定能做到的. 因此 φ 是满射.

单射——由 $A=[T]_{\mathfrak{B}}=0$,当有$[T(b_j)]_{\mathfrak{C}}=0(j=1,2,\cdots,n)$,从而 $T(b_j)=0(j=1,2,\cdots,n)$. 这就是 $T=0$,表明 $\varphi(T)=0$ 蕴含 $T=0$,故 φ 是单射. 定理得证.

注 定理结论可写成 $T\in L(X,Y)\leftrightarrow[T]_{\mathfrak{B}}\in\mathfrak{M}_{m\times n}$；基之间变换公式为 $M_{\mathfrak{C}}=[T]_{\mathfrak{B}}M_{\mathfrak{B}}$；点之间变换公式为 $[T(x)]_{\mathfrak{C}}=[T]_{\mathfrak{B}}[x]_{\mathfrak{B}}$, $\forall x\in X$.

矩阵 $[T]_{\mathfrak{B}}\in\mathfrak{M}_{m\times n}$ 称为**算子** $T\in L(X,Y)$ 的**表示矩阵(representation matrix)**.

上述定理的特例可表述如下.

定理 2.2.2 设 X 是数域$\mathbb{F}$上的有限维线性空间，$\dim X=n$. 则对于 X 的两个基 $\mathfrak{B}$ $\mathfrak{C}$ 算子 $T\in L(X,X)$ 的表示矩阵为 $[T]_{\mathfrak{C}}=M_{\mathfrak{B}}[T]_{\mathfrak{B}}M_{\mathfrak{B}}^{-1}$，其中 $M_{\mathfrak{B}}$ 由 $M_{\mathfrak{B}}\equiv[[b_1]_{\mathfrak{C}}\cdots[b_n]_{\mathfrak{C}}]_{n\times n}$ 确定，$\mathfrak{B}=\left\{e_1=\begin{bmatrix}1\\0\\\vdots\\0\end{bmatrix},\cdots,e_n=\begin{bmatrix}0\\\vdots\\0\\1\end{bmatrix}\right\}$为 X 的标准正交基.

上述定理 2.2.1 还可以推广到三个线性空间之间的情形.

定理 2.2.3 设 X,Y,Z 都是数域$\mathbb{F}$上的有限维线性空间，维数与基分别为 $\dim X=n$，$\mathfrak{B}=\mathfrak{B}$；$\dim Y=m$，$\mathfrak{C}=\mathfrak{C}$；$\dim Z=l$，$\mathfrak{D}=\mathfrak{D}$. 则对于算子 $T\in L(X,Y)$，$S\in L(Y,Z)$，有

$$[S\circ T]_{\mathfrak{B}}=[S]_{\mathfrak{C}}[T]_{\mathfrak{B}},$$

其中 $[T]_{\mathfrak{B}}\in\mathfrak{M}_{m\times n}$，$[S]_{\mathfrak{C}}\in\mathfrak{M}_{l\times m}$，$[S\circ T]_{\mathfrak{B}}\in\mathfrak{M}_{l\times n}$.

证 由

$$x\in X\Rightarrow T(x)\in Y\Rightarrow(S\circ T)(x)=S(T(x))\in Z,$$

得

$$[T(x)]_{\mathfrak{C}}=[T]_{\mathfrak{B}}[x]_{\mathfrak{B}},\quad[S(T(x))]_{\mathfrak{D}}=[S]_{\mathfrak{C}}[T(x)]_{\mathfrak{C}}.$$

于是，将 $[T(x)]_{\mathfrak{C}}=[T]_{\mathfrak{B}}[x]_{\mathfrak{B}}$ 代入后一式，得到

$$[S(T(x))]_{\mathfrak{D}}=[S]_{\mathfrak{C}}[T(x)]_{\mathfrak{C}}=[S]_{\mathfrak{C}}[T]_{\mathfrak{B}}[x]_{\mathfrak{B}},$$

此即

$$[(S\circ T)(x)]_{\mathfrak{D}}=[S]_{\mathfrak{C}}[T]_{\mathfrak{B}}[x]_{\mathfrak{B}},$$

这正是

$$[S\circ T]_{\mathfrak{B}}=[S]_{\mathfrak{C}}[T]_{\mathfrak{B}}.$$

2. 线性算子的像、核、零度

定义 2.2.2(像、核、零度) 设 X,Y 是数域$\mathbb{F}$上的有限维线性空间，对于线性算子空间 $L(X,Y)$ 中的算子 $T\in L(X,Y)$，称 $\mathrm{im}(T)=\{T(x):x\in X\}$ 为 T 的**像(image)**；称 $\ker(T)=\{x\in X:T(x)=0\}$ 为 T 的**核(kernel)**；称 $\dim(\ker(T))$ 为算子 T 的**零度(nullity)**，记为 $\mathrm{null}(T)$；并将像集 $\mathrm{im}(T)$ 的维数称为算子 T 的**秩(rank)**，$\mathrm{rank}(T)=\dim(\mathrm{im}(T))$.

相关的定理列于下面，证明可参看[13].

定理 2.2.4(秩与零度定理) 若 $T\in L(X,Y)$,则 $\dim(\mathrm{im}(T))+\dim(\ker(T))=\dim X$.

定理 2.2.5(第一同构定理) (1) 若 $T\in L(X,Y)$,则商空间 $X/\ker(T)$ 与 $\mathrm{im}(T)$ 同构,即

$$X/\ker(T)\xleftrightarrow{\text{同构}}\mathrm{im}(T);$$

(2) 若 $X_1\subset X$ 是子空间,则商空间 X/X_1 与补子空间 $\complement X_1$ 同构,即

$$X/X_1\xleftrightarrow{\text{同构}}\complement X_1.$$

定理 2.2.6(第二同构定理) 若 $X_1,X_2\subset X$ 是两个子空间,则商空间 $(X_1+X_2)/X_1$ 与商空间 $X_2/(X_1\cap X_2)$ 同构,即

$$(X_1+X_2)/X_1\xleftrightarrow{\text{同构}}X_2/(X_1\cap X_2).$$

定理 2.2.7(第三同构定理) 若 $X_1,X_2\subset X$ 是两个子空间,且 $X_1\subset X_2$,则两个商空间 X/X_1 与 X_1/X_2 的商空间 $(X/X_1)/(X_2/X_1)$ 同构于商空间 X/X_2,即

$$(X/X_1)/((X_2/X_1))\xleftrightarrow{\text{同构}}X/X_2.$$

2.2.2 线性算子的共轭算子

1. 共轭算子的定义

下面介绍与 $T\in L(X,Y)$ 相关的一种重要线性算子.

定义 2.2.3(共轭算子) 设 $T:X\to Y$ 是数域 $\mathbb{F}$ 上的线性空间 X 到 Y 之间的线性映射,$T\in L(X,Y)$. 称映射 $T^*\in L(Y^*,X^*)$ 为 T 的**共轭算子(conjugate operator)**,若对于任意 $x\in X$ 与任意 $f\in Y^*$,成立

$$\langle T^*(f),x\rangle=\langle f,T(x)\rangle,\quad \forall f\in Y^*,\quad \forall x\in X,\tag{2.2.1}$$

这里 $T^*(f)$ 表示 T 对 $f\in Y^*$ 的作用. $T^*\in L(Y^*,X^*)$ 的定义域是 Y^*,值域是 X^*,即

$$T^*:f\in Y^*\xrightarrow{T^*\in L(Y^*,X^*)}T^*(f)\in X^*=L(X,\mathbb{F}).$$

等式(2.2.1)的左边表示 $T^*(f):X\to\mathbb{F}$ 对 $x\in X$ 的作用 $\langle T^*(f),x\rangle\in\mathbb{F}$ 有意义;而等式(2.2.1)的右边 $\langle f,T(x)\rangle,\forall f\in Y^*,\forall x\in X$ 表示 $f\in Y^*=L(Y,\mathbb{F})$ 对 Y 中的元 $y=T(x)\in Y$ 的作用,记为 $\langle f,y\rangle=\langle f,T(x)\rangle,\forall f\in Y^*,y=T(x)\in Y,\forall x\in X$. 于是,$\forall f\in Y^*,\forall x\in X$ 有 $\langle f,T(x)\rangle\in\mathbb{F}$. 图 2.2.1 是(2.2.1)式的共轭算子示意图.

这样,$\forall f\in Y^*,\forall x\in X$,作用 $\langle f,T(x)\rangle\in\mathbb{F}$ 是惟一定义的(well defined). 事实上,如果有一个算子 $S\in L(Y^*,X^*)$ 满足 $\forall f\in Y^*,\forall x\in X$,都成立 $\langle S(f),x\rangle=\langle f,T(x)\rangle$,则此 S 就是算子 $T\in L(X,Y)$ 的**共轭算子**,$S=T^*$,上述等式也成为 $\langle S(f),x\rangle=\langle T^*(f),x\rangle=\langle f,T(x)\rangle,\forall f\in Y^*,\forall x\in X$. 常记(2.2.1)式为 $T^*(f)(x)=f(T(x)),\forall f\in Y^*,\forall x\in X$.

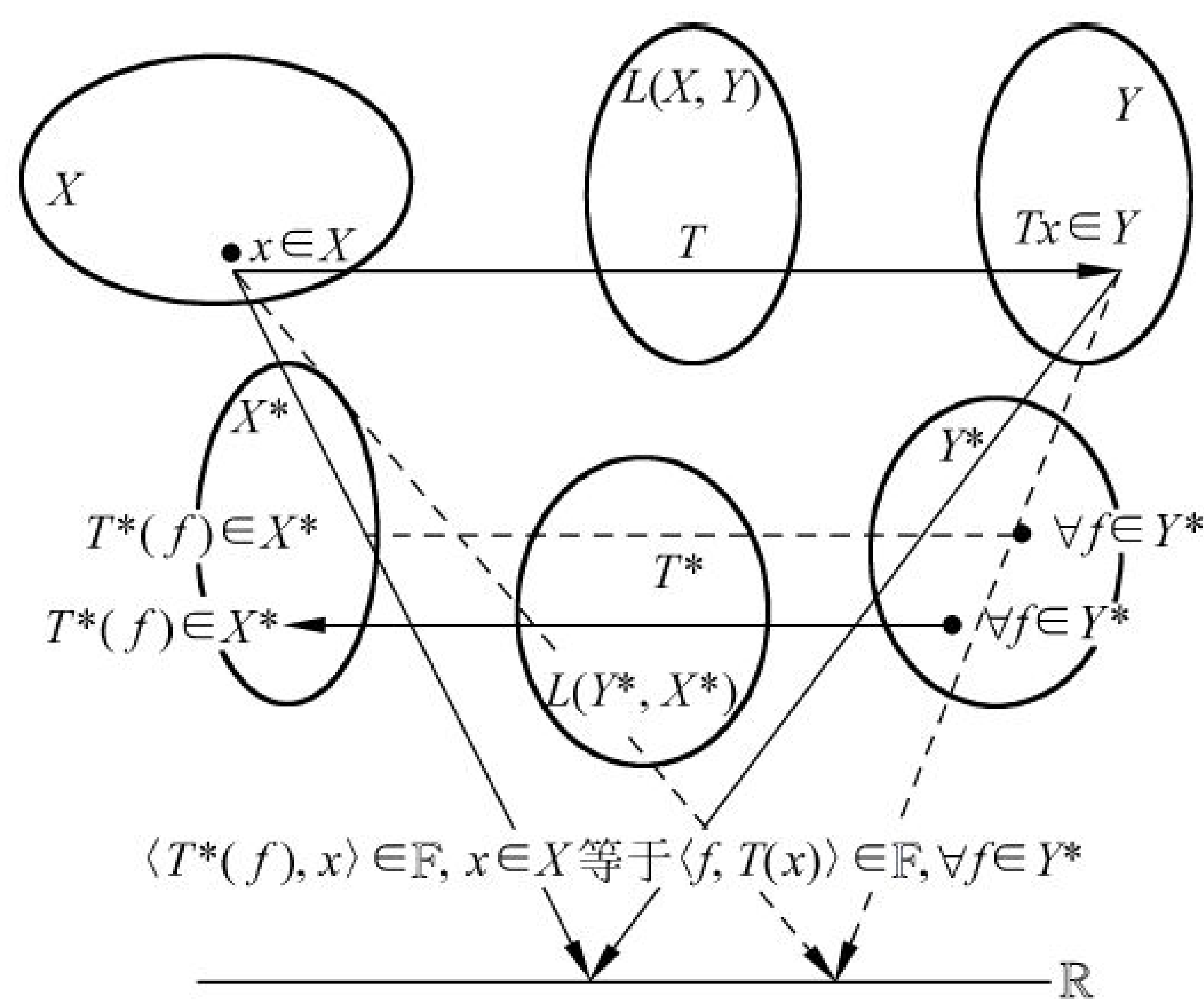


图 2.2.1 共轭算子

定理 2.2.8 设 X,Y 为数域$\mathbb{F}$上的线性空间，则算子 $T,S\in L(X,Y)$的共轭算子 T^*，S^*是线性空间 $L(Y^*,X^*)$中的元，亦即 $T,S\in L(X,Y)\Rightarrow T^*,S^*\in L(Y^*,X^*)$，且

(1) $T,S\in L(X,Y)\Rightarrow (T+S)^*=T^*+S^*$；

(2) $T\in L(X,Y),\alpha\in\mathbb{F}\Rightarrow(\alpha T)^*=\alpha T^*$；

(3) $T\in L(X,Y),U\in L(Y,Z)\Rightarrow(U\circ T)^*=T^*\circ U^*\in L(Z^*,X^*)$；

(4) $T\in L(X,Y)$，若 T^{-1}存在，且 $T^{-1}\in L(Y,X)\Rightarrow(T^{-1})^*=(T^*)^{-1}$；

(5) 若 X,Y 为有限维线性空间，则 $T\in L(X,Y)$的共轭算子 T^*，有$(T^*)^*=T$；

(6) 若 X,Y 为有限维线性空间，$\dim X=n,\dim Y=m$，则当 $T\in L(X,Y)$的表示矩阵为 $A\in\mathfrak{M}_{m\times n}$时，其共轭算子 $T^*\in L(Y^*,X^*)$的表示矩阵为 A 的转置矩阵 $A^{\mathrm{T}}\in\mathfrak{M}_{n\times m}$，且秩满足 $\mathrm{rank}(T)=\mathrm{rank}(T^*)=\mathrm{rank}(A)=\mathrm{rank}(A^{\mathrm{T}})$.

证 依定义，(1)与(2)显然.

对于(3)，任取 $f\in Z^*$，欲证 $\langle(U\circ T)^*(f),x\rangle=\langle f,U(T(x))\rangle$.

由定义，上式左边为$\langle(U\circ T)^*(f),x\rangle=\langle f,(U\circ T)(x)\rangle$，$\forall f\in Z^*$，$\forall x\in X$；于是，再由定义

$$\langle f,(U\circ T)(x)\rangle=\langle f,U(T(x))\rangle=\langle U^*(f),T(x)\rangle=\langle T^*(U^*(f)),x\rangle;$$

上式右边就是$\langle(T^*\circ U^*)(f),x\rangle$，故(3)成立.

对于(4)，取(3)中的$U=T^{-1}$，由假设 $T^{-1}\in L(Y,X)$存在，因此 $I=(I)^*=(T^{-1}\circ T)^*=T^*\circ(T^{-1})^*$，这就给出(4).

对于(5)，由于 X,Y 为有限维的，故 $X\cong X^{**}$，$Y\cong Y^{**}$(定理 2.1.7)，于是

$$T\in L(X,Y)\Rightarrow T^*\in L(Y^*,X^*)\Rightarrow(T^*)^*\in L((X^*)^*,(Y^*)^*)=L(X,Y).$$

我们还需要证明，$\forall x\in X$，有$\langle T,x\rangle=\langle T^{**},x\rangle$.

其实,对任意 $x^{**}\in X^{**}$,对应地 $x^{**}\in X^{**}\leftrightarrow x\in X$,且此对应关系由

$$\langle f,x\rangle=\langle x^{**},f\rangle,\quad \forall f\in Y^{*} \tag{2.2.2}$$

确定,因此,对于 $T^{**}\in L(X^{**},Y^{**})$以及任意 $f\in Y^{*}$,有

$$\langle T^{**}(x^{**}),f\rangle \xlongequal{\text{定义}} \langle x^{**},T^{*}(f)\rangle \xlongequal{\text{定义}} \langle x^{**},f(T)\rangle \xlongequal{(2.2.2)\text{式}} \langle f(T),x\rangle \xlongequal{\text{定义}} \langle T(x),f\rangle,$$

从而,$T^{**}(x^{**})=T(x)$. 但 $X\cong X^{**}$,$Y\cong Y^{**}$,因此 $T^{**}(x^{**})=T(x)$表示 T^{**} 与 T 将 $\forall x^{**}\in X^{**}\leftrightarrow x\in X$ 映为同一个值,所以 $T^{**}=T$.

最后,对(6),由于 X,Y 为有限维的,设 $\mathfrak{B}=\{b_1,b_2,\cdots,b_n\}$,$\mathfrak{C}=\{c_1,c_2,\cdots,c_m\}$. 对偶基为 $\mathfrak{B}_{*}=\{b_1^{*},b_2^{*},\cdots,b_n^{*}\}$,$\mathfrak{C}_{*}=\{c_1^{*},c_2^{*},\cdots,c_m^{*}\}$. 于是,算子 $T\in L(X,Y)\leftrightarrow[T]_{\mathfrak{B}}\in\mathfrak{M}_{m\times n}$. 另一方面,$T^{*}\in L(Y^{*},X^{*})\leftrightarrow[T^{*}]_{\mathfrak{C}^{*}\mathfrak{B}^{*}}\in\mathfrak{M}_{n\times m}$.

由 2.2.1 节的讨论知

$$[T]_{\mathfrak{B}}=([T(b_1)]_{\mathfrak{C}}\ [T(b_2)]_{\mathfrak{C}}\cdots\ [T(b_n)]_{\mathfrak{C}})$$

与

$$[T^{*}]_{\mathfrak{C}^{*}\mathfrak{B}^{*}}=([T^{*}(c_1^{*})]_{\mathfrak{B}^{*}}\ \ [T^{*}(c_2^{*})]_{\mathfrak{B}^{*}}\ \cdots\ [T^{*}(c_m^{*})]_{\mathfrak{B}^{*}}),$$

直接用坐标表示,便可比较出两个矩阵的关系(请读者自行验证),从而得到$[T^{*}]_{\mathfrak{C}^{*}\mathfrak{B}^{*}}=([T]_{\mathfrak{B}})^{\mathrm{T}}$. 定理得证.

例 2.2.1 设 X,Y 为 $\mathbb{R}$ 上的线性空间,$\dim X=n$,$\dim Y=m$. 则

$$\dim X^{*}=n,\quad \dim Y^{*}=m,$$

并且 $L(X,Y)\leftrightarrow\mathfrak{M}_{m\times n}$,$L(Y^{*},X^{*})\leftrightarrow\mathfrak{M}_{n\times m}$.

由定理 2.2.8,若 $T\in L(X,Y)$与 $T^{*}\in L(Y^{*},X^{*})$的表示矩阵分别为 $A_{m\times n}$与 $B_{n\times m}$,即 $T\in L(X,Y)\leftrightarrow A_{m\times n}\in\mathfrak{M}_{m\times n}$,$T^{*}\in L(Y^{*},X^{*})\leftrightarrow B_{n\times m}\in\mathfrak{M}_{n\times m}$,则 $B_{n\times m}=(A_{m\times n})^{\mathrm{T}}$ 是 $A_{m\times n}$的转置矩阵.

2. 内积空间上的共轭算子的两种等价定义

本小节中,假设 X,Y,Z 等都是内积空间.

1) Riesz 表现定理

Riesz 表现定理在内积空间 X 的对偶空间 X^{*} 中,有如下形式(与定理 2.1.8 比较,并获得其证明).

定理 2.2.9 设$(X,(x,y))$是内积空间,(x,y)是 X 上的内积,X^{*} 是其对偶空间,则任一 $f\in X^{*}$,存在惟一的元 $x\in X$ 与 f 对应,$f\leftrightarrow x$,使得对任意 $v\in X$,线性泛函 f 对 $v\in X$ 的"作用"可表示为 $f(v)\equiv(v,x)$,其中(v,x)是 X 上的内积.

2) 共轭算子的两种定义

我们考虑"共轭算子"在内积空间中的另一种等价定义.

对于内积空间 X 上的线性泛函 $f\in X^{*}=L(X,\mathbb{F})$,由 Riesz 表现定理 2.2.9,有

$$\forall f\in X^{*}\Rightarrow\exists!x\in X,\ \mathrm{s.t.}\ f(v)=(v,x),\forall v\in X,$$

因此，$f\in X^*\leftrightarrow x\in X$ 是一对一的，记为 $J:f\in X^*\to x\in X$，$J(f)=x$．于是，代入 $f(v)=(x,v)$，得 $f(v)=(v,J(f))$．

$J(f)$有如下性质：$\forall v\in X$，$f,g\in X^*$，$\alpha,\beta\in\mathbb{F}$，则

$$\begin{aligned}(v,J(\alpha f+\beta g))&=(\alpha f+\beta g)(v)=\alpha f(v)+\beta g(v)=(\alpha f)(v)+(\beta g)(v)\\&=(v,\bar{\alpha}J(f))+(v,\bar{\beta}J(g))=(v,\bar{\alpha}J(f)+\bar{\beta}J(g)),\end{aligned}$$

由此得到 $J(\alpha f+\beta g)=\bar{\alpha}J(f)+\bar{\beta}J(g)$，所以，$J:X^*\to X$ 是线性的（当 X 是实内积空间）、或共轭线性的（当 X 是复内积空间），并且它也是满射与单射，因此，$J:X^*\to X$ 是 X^* 到 X 的"共轭同构"，这里，共轭同构在实内积空间情形，就是同构；在复内积情形，是共轭同构，统称共轭同构．

于是，由内积空间$(X,(x,y))$上的内积(x,y)出发，可以定义 X^* 上的内积

$$(f,g)=(J^{-1}f,J^{-1}g),\quad f,g\in X^*,\tag{2.2.3}$$

不难验证，$(X^*,(f,g))$构成内积空间．

由 $X^*\xleftrightarrow{J_X}X$，$Y^*\xleftrightarrow{J_Y}Y$，因此 $L(Y^*,X^*)\leftrightarrow L(Y,X)$；于是，在共轭算子的定义(2.2.1)式中，在同构的意义下，可将 $T^*\in L(Y^*,X^*)$换为 $T'\in L(Y,X)$，将 $\forall f\in Y^*$ 换为 $\forall y\in Y$，从而得到用内积$(T'y,x)=(y,Tx)$，$\forall y\in Y$，$\forall x\in X$ 表示的等式．

定义 2.2.4（共轭算子）　设 $T:X\to Y$ 是数域$\mathbb{F}$ 上的内积空间 X 到内积空间 Y 之间的映射．对于 $T\in L(X,Y)$，称 $T'\in L(Y,X)$为 T 的**共轭算子**，若

$$(T'y,x)=(y,Tx),\quad \forall y\in Y,\forall x\in X,\tag{2.2.4}$$

等式(2.2.4)左边$(T'y,x)$是 X 的内积，右边(y,Tx)是 Y 的内积．

3）共轭算子两种定义的关系

设 $T^*:Y^*\to X^*$ 与 $T':Y\to X$ 为内积空间的共轭算子的两种定义，分别由(2.2.1)式、(2.2.4)式给出．若 $\varphi_1:X^*\leftrightarrow X$，$\varphi_2:Y^*\leftrightarrow Y$ 为同构映射，则定义 $\sigma:Y^*\to X^*$ 为（图 2.2.2 中虚线）$\sigma=(\varphi_1)^{-1}T'\varphi_2$，这里 $\sigma=(\varphi_1)^{-1}T'\varphi_2:Y^*\to X^*$ 是线性映射，并且 $\forall x\in X$，若$(\varphi_1)^{-1}(x)=x^*\in X^*$，则 $\varphi_1(x^*)=x$，所以

$$((\varphi_1)^{-1}(x),v)=(x^*,v)=(v,\varphi_1(x^*)),\quad \forall v\in X.$$

于是，$\forall f\in Y^*$，$\forall v\in X$，有

$$\begin{aligned}(\sigma(f),v)&=[(\varphi_1)^{-1}T'\varphi_2(f)](v)\\&=(\varphi_1)^{-1}[T'\varphi_2(f)](v)\\&=(v,T'\varphi_2(f))=(T(v),\varphi_2(f))\\&=f(T(v))=(T^*(f),v).\end{aligned}$$

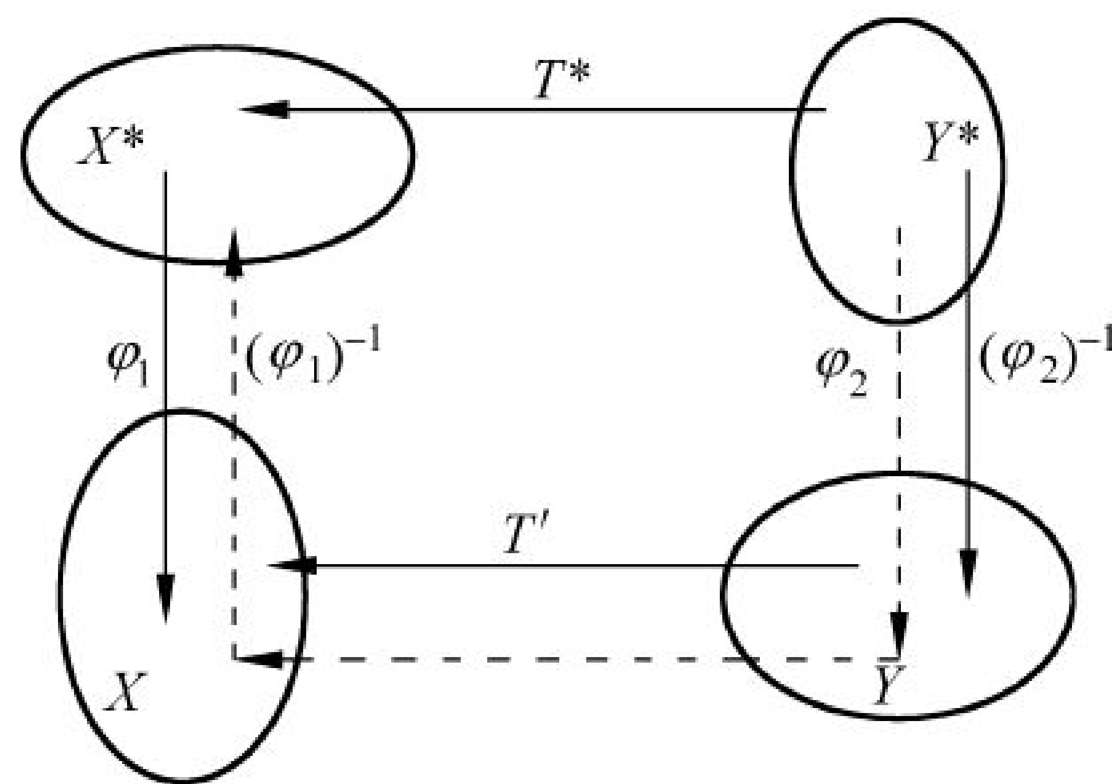


图 2.2.2　T^* 与 T'

从而得 $\sigma=T^*$，即 $T^*=(\varphi_1)^{-1}T'\varphi_2$．于是，$T^*$ 与 T 是彼此同构的．图 2.2.2 是一个交换图．

通常，在有限维内积空间中多取共轭算子为 T'，因为定义 2.2.4 是用内积表示的，更为

方便.

若 $\mathfrak{B}=\{b_1,b_2,\cdots,b_n\}$ 与 $\mathfrak{C}=\{c_1,c_2,\cdots,c_m\}$ 分别是 X 与 Y 的标准正交基,并设 $\forall T\in L(X,Y)$,有

$$[T]_{\mathfrak{B}}=[[T(b_1)]_{\mathfrak{C}}\cdots[T(b_n)]_{\mathfrak{C}}]\equiv[\beta_{ij}]_{m\times n},$$

其中 $\beta_{ij}=(T(b_i),c_j)$. 另一方面,$T'\in L(Y,X)$,则

$$[T']_{\mathfrak{C}}=[[T'(c_1)]_{\mathfrak{B}}\cdots[T'(c_m)]_{\mathfrak{B}}]\equiv[\gamma_{ij}]_{n\times m},$$

其中 $\gamma_{ij}=(T'(c_i),b_j)$. 注意到 X,Y 均为内积空间,故

$$\gamma_{ij}=(T'(c_i),b_j)=\overline{(b_j,T'(c_i))}=\overline{(T(b_j),c_i)}=\overline{\beta_{ji}},\quad i=1,2,\cdots,m,\quad j=1,2,\cdots,n.$$

因此

$$[T']_{\mathfrak{C}}=\overline{([T]_{\mathfrak{B}})}^{\mathrm{T}},$$

这样,我们得到:在有限维内积空间的中,若算子 $T:X\to Y$ 的表示矩阵为$[T]_{\mathfrak{B}}=[a_{jk}]_{m\times n}$,则其共轭算子 $T':Y\to X$ 的表示矩阵为矩阵$[T]_{\mathfrak{B}}$的共轭转置,即为$\overline{([T]_{\mathfrak{B}})}^{\mathrm{T}}=\overline{[a_{kj}]}_{n\times m}$. 因此,对于内积空间,其共轭算子有如下定理.

定理 2.2.10 设 X,Y 为数域$\mathbb{F}$上的内积空间,则共轭算子 $T'\in L(Y,X)$在定义 2.2.4 之下,有

(1) $T,S\in L(X,Y)\Rightarrow(T+S)'=T'+S'$;

(2) $T\in L(X,Y),\alpha\in\mathbb{F}\Rightarrow(\alpha T)'=\bar{\alpha}T'$;

(3) $T\in L(X,Y),U\in L(Y,Z)\Rightarrow(U\circ T)'=T'\circ U'$;

(4) $T\in L(X,Y)$,且 T^{-1}存在 $\Rightarrow(T^{-1})'=(T')^{-1}$;

(5) 若 X,Y 为有限维线性空间,则对 $T\in L(X,Y)$的共轭算子 T',有$(T')'=T$;

(6) 若 X,Y 为有限维线性空间,$\dim X=n,\dim Y=m$,则当 $T\in L(X,Y)$的表示矩阵为 $A\in\mathfrak{M}_{m\times n}$时,其共轭算子 $T'\in L(Y,X)$的表示矩阵 B 为 $A\in\mathfrak{M}_{m\times n}$的共轭转置矩阵 $B=(\bar{A})^{\mathrm{T}}\in\mathfrak{M}_{n\times m}$,且秩 $\mathrm{rank}(T)=\mathrm{rank}(T')=\mathrm{rank}(A)=\mathrm{rank}(B)$.

3. 内积空间中的特殊共轭算子

定义 2.2.5(特殊的共轭算子) 设$(X,+,\alpha\cdot,(x,y))$是数域$\mathbb{F}$上的内积空间,算子 $T\in L(X,X)$,其共轭算子为 $T'\in L(X,X)$,由定义 2.2.4 给出(见(2.2.4)式):

$$(T'y,x)=(y,Tx),\quad\forall y\in X,\forall x\in X.$$

(1) 若 $T=T'$,则称 T 为**自共轭(self-conjugate)算子**,或**埃尔米特(Hermite)算子**;

(2) 若 T 为一一映射,且 $T'=T^{-1}$,则称 T 为**酉算子(unitary operator)**;

(3) 若 $TT'=T'T$,则称 T 为**正规算子(normal operator)**.

相应地,设 $T\in L(X,X)\leftrightarrow A\in\mathfrak{M}_n$,$T'\in L(X,X)\leftrightarrow B\in\mathfrak{M}_n$.

当 $A\in\mathfrak{M}_n(\mathbb{C})$为复方阵时,$B$ 为 A 的共轭转置方阵,$B=\bar{A}^{\mathrm{T}}$,有

(1)′ 若 $B=A$,则称 A 为**埃尔米特方阵**(或**自共轭方阵**);

(2)′ 若 $B=-A$,则称 A 为**斜埃尔米特方阵**;

(3)′ 若 A 为可逆方阵,且 $B=A^{-1}$,则称 A 为**酉方阵**;

(4)′ 若 $AB=BA$,则称 A 为**正规方阵**;

当 $A\in\mathfrak{M}_n(\mathbb{R})$为实方阵时,$B$ 为 A 的转置方阵,$B=A^{\mathrm{T}}$,有

(1)″ 若 $A^{\mathrm{T}}=A$,则称 A 为**对称方阵**;

(2)″ 若 $A^{\mathrm{T}}=-A$,则称 A 为**斜对称方阵**;

(3)″ 若 A 为可逆方阵,且 $A^{\mathrm{T}}=A^{-1}$,则称 A 为**正交方阵**.

易见,$\mathbb{F}=\mathbb{R}$ 时,自共轭方阵就是对称方阵 $A=B$;酉方阵就是正交方阵 $A^{-1}=B$.

例 2.2.2 埃尔米特算子与埃尔米特方阵在物理学中有重要应用,我们给出几个特殊的例子.在二级能量系统、角动量问题中会遇到的方阵有以下几个:

$$I=\begin{bmatrix}1&0\\0&1\end{bmatrix},\quad I^{\mathrm{T}}=\begin{bmatrix}1&0\\0&1\end{bmatrix}^{\mathrm{T}}=\begin{bmatrix}1&0\\0&1\end{bmatrix}=I\text{ 是埃尔米特方阵(对称方阵)};$$

$$A=\begin{bmatrix}0&1\\1&0\end{bmatrix},\quad A^{\mathrm{T}}=\begin{bmatrix}0&1\\1&0\end{bmatrix}^{\mathrm{T}}=\begin{bmatrix}0&1\\1&0\end{bmatrix}=A\text{ 是埃尔米特方阵(对称方阵)};$$

$$B=\begin{bmatrix}1&0\\0&-1\end{bmatrix},\quad B^{\mathrm{T}}=\begin{bmatrix}1&0\\0&-1\end{bmatrix}^{\mathrm{T}}=\begin{bmatrix}1&0\\0&-1\end{bmatrix}=B\text{ 是对称方阵,称为泡利方阵};$$

$$C=\begin{bmatrix}0&1\\-1&0\end{bmatrix},\quad C^{\mathrm{T}}=\begin{bmatrix}0&1\\-1&0\end{bmatrix}^{\mathrm{T}}=\begin{bmatrix}0&-1\\1&0\end{bmatrix}=-C\text{ 是斜对称方阵};$$

$$D=\begin{bmatrix}0&\mathrm{i}\\\mathrm{i}&0\end{bmatrix},\quad D=\overline{\begin{bmatrix}0&\mathrm{i}\\\mathrm{i}&0\end{bmatrix}}^{\mathrm{T}}=\begin{bmatrix}0&-\mathrm{i}\\-\mathrm{i}&0\end{bmatrix}=-\begin{bmatrix}0&\mathrm{i}\\\mathrm{i}&0\end{bmatrix}=-D\text{ 是斜埃尔米特方阵};$$

$$E=\begin{bmatrix}0&-\mathrm{i}\\\mathrm{i}&0\end{bmatrix},\quad E=\overline{\begin{bmatrix}0&-\mathrm{i}\\\mathrm{i}&0\end{bmatrix}}^{\mathrm{T}}=\begin{bmatrix}0&-\mathrm{i}\\\mathrm{i}&0\end{bmatrix}=E\text{ 是埃尔米特方阵.}$$

请读者指出这些方阵中哪些是酉方阵. 三维方阵

$$\begin{bmatrix}0&1&0\\1&0&1\\0&1&0\end{bmatrix},\quad\begin{bmatrix}0&-\mathrm{i}&0\\\mathrm{i}&0&-\mathrm{i}\\0&\mathrm{i}&0\end{bmatrix},\quad\begin{bmatrix}1&0&0\\0&0&0\\0&0&-1\end{bmatrix}$$

在三维角动量问题中更起着重要作用. 请读者自行验证它们的类型,指出哪些是酉方阵.

我们分别列出三类重要的算子的重要性质.

定理 2.2.11 设$(X,+,\alpha\cdot,(x,y))$为数域 $\mathbb{F}$ 上的内积空间,设 $T,S\in L(X,X)$,则有下列结论:

(1) 若 T 是自共轭算子,则 $\forall x,y\in X$,有 $((T,y),x)=(x,(T,y))$;

(2) 若 T,S 是自共轭的,则 $T+S$ 也是自共轭的;

(3) 若 T 是自共轭的,则对于实数 $\alpha\in\mathbb{R}$,αT 也是自共轭的;

(4) 若 T 是自共轭的，且 T 可逆，则 T^{-1} 也是自共轭的；

(5) 若 T 是自共轭的，则 $\forall x\in X$，内积 $((T,x),x)\in\mathbb{R}$ 为实值；

(6) 若 T 是自共轭的，则 $\forall x\in X$，$((T,x),x)=0$，蕴含 $T=0$.

定理 2.2.12 设 $(X,+,\alpha\cdot,(x,y))$ 为数域 $\mathbb{F}$ 上的内积空间，设 $T,S\in L(X,X)$，则有下列结论：

(1) 若 T 是酉算子，则 $\forall x,y\in X$，有 $((T,x),y)=(x,(T^{-1},y))$；

(2) 若 T 是酉算子，则 T^{-1} 也是酉算子；

(3) 若 T,S 是酉算子，则 $T\circ S$ 也是酉算子；

(4) 若 T 是酉算子，当且仅当 T 是满射且保内积的，亦即，

$$\forall x,y\in X \Rightarrow ((T,x),(T,y))=(x,y);$$

(5) 若 $\dim X=n$ 为有限维的，则 T 是酉算子，当且仅当 T 将 X 的标准正交基映为标准正交基.

对应地，关于表示方阵，若 $T\in L(X,X)\leftrightarrow A\in\mathfrak{M}_n$，我们有

(1)′ A 是酉方阵，当且仅当 A 的所有列向量组成 $\mathbb{C}^n$ 的一组标准正交基，也当且仅当 A 的所有行向量组成 $\mathbb{C}^n$ 的一组标准正交基；

(2)′ 若 A 是酉方阵，则 $|\det(A)|=1$；

(3)′ 若 A 是正交方阵，则 $\det(A)=\pm1$.

定理 2.2.13 设 $(X,+,\alpha\cdot,(x,y))$ 为数域 $\mathbb{F}$ 上的内积空间，设 $T\in L(X,X)$ 为正规算子，则

(1) 对于 $x\in X$，$T(x)=0$ 蕴含 $T'(x)=0$；

(2) $\forall k\in\mathbb{N}$，$k>0$，$T^k(x)=0$ 蕴含 $T(x)=0$；

(3) 若 $\forall x\in X$，$T(x)=\lambda x$，$\lambda\in\mathbb{F}$，则 $T'(x)=\bar{\lambda}x$.

4. 代数、算子代数

在线性空间 $(X,+,\alpha\cdot)$ 中，常常可以引进元之间的乘法运算 $\times$，形成一种新的运算结构，称为代数.

定义 2.2.6(代数) 设 $(X,+,\alpha\cdot)$ 是数域 $\mathbb{F}$ 上的线性空间，在 X 中定义元之间的运算"乘法"：$x\in X$ 与 $y\in X$ 的乘法运算 $x\times y$，简记为 xy，使满足

(1) $x,y\in X \Rightarrow x\times y\in X$； (封闭性)

(2) $x,y,z\in X \Rightarrow (x\times y)\times z=x\times(y\times z)$； (结合律)

(3) $x,y,z\in X \Rightarrow x\times(y+z)=x\times y+x\times z,(y+z)\times x=y\times x+z\times x$；(加乘分配律)

(4) $x,y\in X$，$\alpha\in\mathbb{F}\Rightarrow \alpha\cdot(x\times y)=(\alpha\cdot x)\times y=x\times(\alpha\cdot y)$， (数乘结合律)

则称 $(X,+,\alpha\cdot,\times)$ 为数域 $\mathbb{F}$ 上的一个代数(algebra)，简称**代数**.

注意,(1)与(2)表示 X 关于乘法运算构成一个半群.

进而,若还满足

(5) 存在运算×的单位元 $I\in X$,使得 $\forall x\in X$,有 $I\times x=x\times I$, (乘法单位元)

则称 $(X,+,\alpha\cdot,\times)$ 为数域 $\mathbb{F}$ 上的一个具有单位元的代数(algebra with unit).

再进一步,若还满足

(6) $x,y\in X \Rightarrow x\times y=y\times x$, (乘法交换律)

则称 $(X,+,\alpha\cdot,\times)$ 为数域 $\mathbb{F}$ 上的一个具有单位元的交换代数(exchange algebra with unit).

例 2.2.3 $C([a,b])$ 为区间 $[a,b]$ 上连续函数的集合

$C([a,b])$ 中的加法、数乘、乘法分别定义为

$$f,g\in C([a,b])\Rightarrow (f+g)(x)=f(x)+g(x),\quad x\in[a,b];$$

$$f,g\in C([a,b])\Rightarrow (\alpha f)(x)=\alpha f(x),\quad \alpha\in\mathbb{R},\quad x\in[a,b];$$

$$f,g\in C([a,b])\Rightarrow (f\times g)(x)=f(x)g(x),\quad x\in[a,b].$$

则 $(C([a,b]),+,\alpha\cdot,\times)$ 成为一个具有单位元 $I(x)\equiv 1\in C([a,b])$ 的交换代数.

例 2.2.4 $L^1([a,b])$ 为区间 $[a,b]$ 上 Lebesgue 可积函数的集合

$L^1([a,b])$ 中的加法、数乘与 $C([a,b])$ 中相同. 乘法定义为

$$f,g\in L^1([a,b])\Rightarrow (f*g)(x)=\int_{[a,b]} f(x-t)g(t)\,\mathrm{d}t,$$

其中运算 $*$ 称为**卷积**,则 $(L^1([a,b]),+,\alpha\cdot,*)$ 成为一个没有单位元的交换代数.

代数 $(L^1([a,b]),+,\alpha\cdot,*)$ 没有单位元,在今后介绍 Fourier 分析时,将证明这一点.

例 2.2.5 线性算子空间 $(L(X,X),+,\alpha\cdot)$

定义算子的加法与数乘为

$$T,S\in L(X,X)\Rightarrow (T+S)(x)=T(x)+S(x),\quad x\in X;$$

$$T\in L(X,X),\alpha\in\mathbb{F}\Rightarrow (\alpha T)(x)=\alpha T(x),\quad x\in X,$$

乘法定义为复合运算 $\circ$,即

$$T,S\in L(X,X)\Rightarrow (S\circ T)(x)=S(T(x)),\quad x\in X,$$

则 $(L(X,X),+,\alpha\cdot,\circ)$ 成为一个具有单位元 $I\in L(X,X)$ 的不可交换代数,$I(x)=x$, $\forall x\in X$,称为 X 上的**算子代数**(**operator algebra**).

2.2.3 多重线性代数

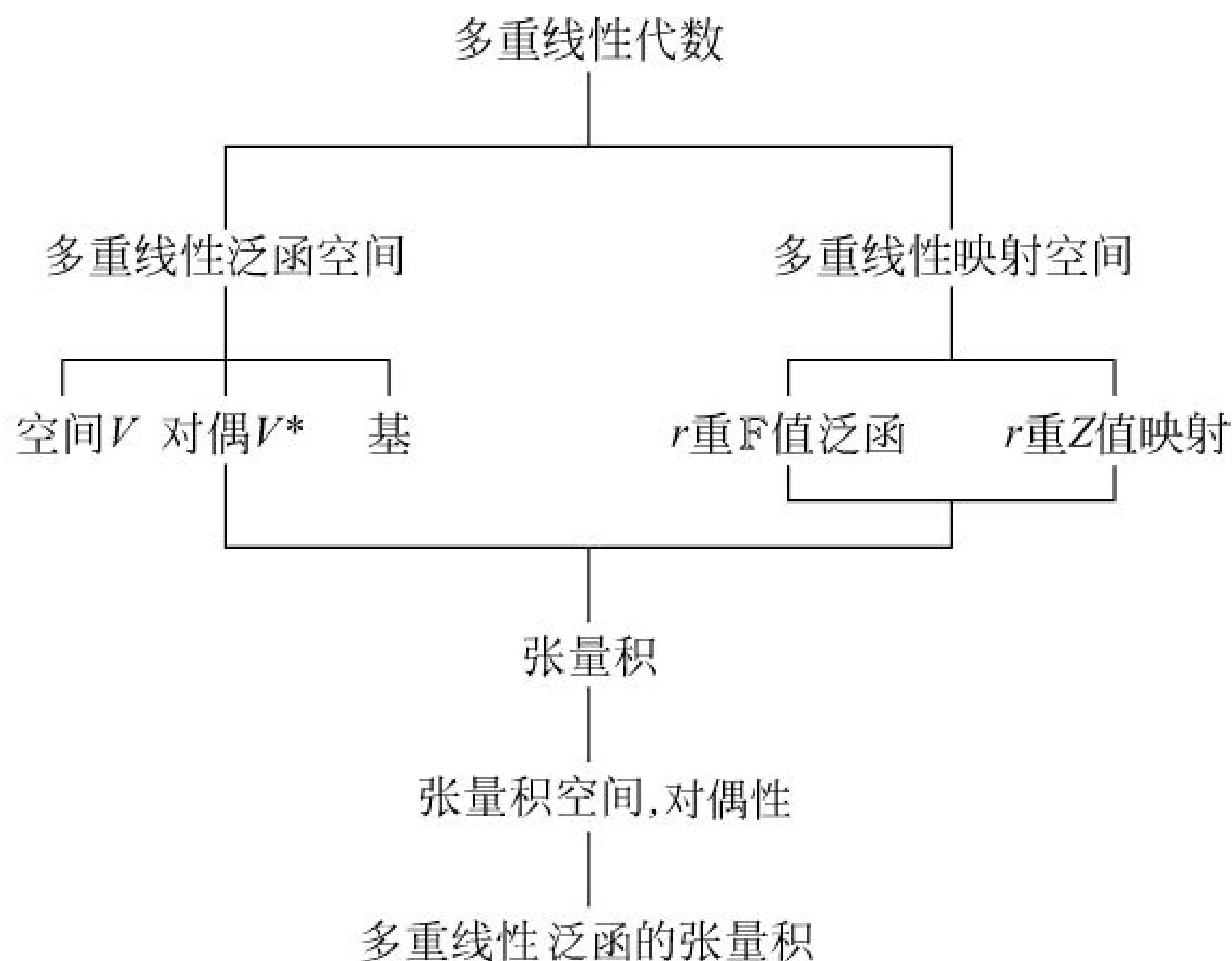
多重线性代数是线性代数的重要组成部分之一,其应用非常广泛. 我们主要介绍多重线性映射空间

$$L(V_1,V_2,\cdots,V_r;Z)\equiv L(V_1\times V_2\times\cdots\times V_r,Z)$$

与多重线性泛函空间

$$L(V_1,V_2,\cdots,V_r;\mathbb{F})\equiv L(V_1\times V_2\times\cdots\times V_r,\mathbb{F}),$$

其中 $V_1,V_2,\cdots,V_r$ 是同一数域$\mathbb{F}$上的 r 个线性空间.进而研究它们的对偶空间,以及其间的关系. 以这些为基础,进一步研究张量积、张量积空间等,为微分几何的学习打下基础.



1. 空间 V,V^*,基与对偶基

本节考虑数域$\mathbb{F}$上的线性空间 $V,W,\cdots$. 记它们的对偶空间(即其上的线性泛函空间)为 $V^*,W^*,\cdots$.

设 $\mathfrak{A}=\{a_1,a_2,\cdots,a_n\}$是 n 维线性空间 V 的一个基,且 $\forall v\in V$,有 $v=\sum\limits_{j=1}^{n}v_ja_j$;记 $[v]_{\mathfrak{A}}=\begin{bmatrix}v_1\\v_2\\\vdots\\v_n\end{bmatrix}$为元 $v\in V$ 关于基 $\mathfrak{A}=\{a_1,a_2,\cdots,a_n\}$的坐标,$v_j\in\mathbb{F}\ (1\leqslant j\leqslant n)$. 于是,对于 $f\in V^*$,有

$$f(v)=\sum_{j=1}^{n}v_jf(a_j)\equiv\sum_{j=1}^{n}v_jf_j. \tag{2.2.5}$$

所以线性泛函 $f\in V^*$ 由它在基 $\mathfrak{A}=\{a_1,a_2,\cdots,a_n\}$上的值$\begin{bmatrix}f(a_1)\\f(a_2)\\\vdots\\f(a_n)\end{bmatrix}\equiv\begin{bmatrix}f_1\\f_2\\\vdots\\f_n\end{bmatrix}$确定. 在 V 的对偶空间 V^* 中取一组线性泛函$\{a_1^*,a_2^*,\cdots,a_n^*\}$,满足

$$a_j^*(a_k)=\delta_{kj},\quad 1\leqslant j,k\leqslant n, \tag{2.2.6}$$

其中 $\delta_{kj}=\begin{cases}1, & j=k,\\0, & j\neq k.\end{cases}$因此

$$a_j^*(v) = a_j^*\left(\sum_{k=1}^n v_k a_k\right) = v_j, \quad 1 \leqslant j \leqslant n.$$

从而对于(2.2.5)式，由 $v_j = a_j^*(v)$，得

$$\forall v \in V, v = \sum_{j=1}^n v_j a_j \Rightarrow f(v) = \sum_{j=1}^n v_j f(a_j) = \sum_{j=1}^n a_j^*(v) f(a_j) = \sum_{j=1}^n f_j a_j^*(v).$$

这样，对于线性泛函 $f \in V^*$，就有表现形式 $f = \sum_{j=1}^n f_j a_j^*$，这里 $f_j = f(a_j)$，$1 \leqslant j \leqslant n$，且这种表示是惟一的，故 $\{a_j^*, 1 \leqslant j \leqslant n\}$ 构成 V^* 的基，称为 V 的基 $\mathfrak{A} = \{a_j, 1 \leqslant j \leqslant n\}$ 的**对偶基**(**dual basis**)，记为 $\mathfrak{A}^* = \{a_j^*, 1 \leqslant j \leqslant n\}$.

定理 2.2.14 对于有限维线性空间 V，$\dim V = n$，则其对偶空间 V^* 是数域 $\mathbb{F}$ 上的线性空间，且有

(1) $\dim V = \dim V^* = n$；

(2) 若 $\mathfrak{A} = \{a_j, 1 \leqslant j \leqslant n\}$ 是 V 的一个基，则 $\mathfrak{A}^* = \{a_j^*, 1 \leqslant j \leqslant n\}$ 是 V^* 中的对偶基，满足 $a_j^*(a_k) = \delta_{kj}$ $(1 \leqslant j, k \leqslant n)$；

(3) $(V^*)^* = V$.

2. 多重线性映射空间

对于有限维线性空间 V 与 V^*，$\dim V = \dim V^* = n$，有如下定义.

定义 2.2.7($V \times V^*$ **上的**($\mathbb{F}$ **值**)**线性泛函**) 设 V，$\dim V = n$，是数域 $\mathbb{F}$ 上的线性空间，V^* 是其对偶空间，对于每一 $v^* \in V^*$，定义

$$v^*(v) = \langle v^*, v \rangle, \quad \forall v \in V,$$

这里的 $\langle v^*, v \rangle \in \mathbb{F}$ 是定义在 $V \times V^*$ 上、取值于 $\mathbb{F}$ 中的函数，并且关于变量 v，泛函 v^* 都是线性的，从而是双线性的，亦即

$$\begin{cases} \langle v^*, \alpha_1 v_1 + \alpha_2 v_2 \rangle = \alpha_1 \langle v^*, v_1 \rangle + \alpha_2 \langle v^*, v_2 \rangle, \\ \langle \alpha_1 v_1^* + \alpha_2 v_2^*, v \rangle = \alpha_1 \langle v_1^*, v \rangle + \alpha_2 \langle v_2^*, v \rangle, \end{cases} \quad \alpha_1, \alpha_2 \in \mathbb{F},$$

其中 $v, v_1, v_2 \in V$，$v^*, v_1^*, v_2^* \in V^*$，则称 $\langle v^*, v \rangle : V^* \times V \to \mathbb{F}$ 为 $V^* \times V$ 上的 $\mathbb{F}$ **值线性泛函**.

注 1 通常将定义中的 $\langle v^*, v \rangle$ 与 $\langle v, v^* \rangle$ 视为一致，$\langle v, v^* \rangle \equiv \langle v^*, v \rangle$，这在有限维线性空间中是完全合理的. 因为 $\langle v, v^* \rangle$ 是将 $v \in V$ 视为 V^* 上的线性泛函，而 $\langle v^*, v \rangle$ 是将 v^* 视为 V 上的线性泛函，但因 $V \leftrightarrow V^* \leftrightarrow V^{**}$，故当 v 是 V^* 上的线性泛函时，对应的 v^* 也就成为 V 上的线性泛函，且二者的值是相同的.

注 2 定义中的 $\langle v, \cdot \rangle$ 是 V^* 上的 $\mathbb{F}$ 值线性泛函；反之，V^* 上的任一个 $\mathbb{F}$ 值线性泛函 $\varphi : V^* \to \mathbb{F}$ 都可表示为 $\langle v, \cdot \rangle$ 的形式. 因为对于给定的 $\varphi : V^* \to \mathbb{F}$，令 $v = \sum_{j=1}^n \varphi(a_j^*) a_j$，则

对任一 $v^* \in V^*$，都有 $\langle v, v^* \rangle = \varphi(v^*)$. 这也又一次表明 V 也是 V^* 的对偶空间.

注 3 定义 2.2.7 中的$\mathbb{F}$值线性泛函可以推广到"多重线性泛函"上去.

定义 2.2.8（r **重（$\mathbb{F}$ 值）线性泛函、r 重（Z 值）线性映射**）

（1）**二重（$\mathbb{F}$ 值）线性泛函** 设 V, W 是$\mathbb{F}$上的线性空间，若泛函 $f: V \times W \to \mathbb{F}$ 对于每个变量都是线性的，亦即

$$\begin{cases} f(\alpha_1 v_1 + \alpha_2 v_2, w) = \alpha_1 f(v_1, w) + \alpha_2 f(v_2, w), \\ f(v, \alpha_1 w_1 + \alpha_2 w_2) = \alpha_1 f(v, w_1) + \alpha_2 f(v, w_2), \end{cases} \quad \alpha_1, \alpha_2 \in \mathbb{F},$$

其中 $v, v_1, v_2 \in V, w, w_1, w_2 \in W$，则称 f 为积空间 $V \times W$ 上的**二重（$\mathbb{F}$ 值）线性泛函**，记这样的线性泛函（注意上式为$\mathbb{F}$中的数的等式）的全体为 $L(V, W; \mathbb{F}) \equiv L(V \times W, \mathbb{F})$，它在以下运算下构成$\mathbb{F}$上的线性空间：

$$f, g \in L(V, W; \mathbb{F}) \Rightarrow (f+g)(v, w) = f(v, w) + g(v, w);$$

$$f \in L(V, W; \mathbb{F}), \quad \alpha \in \mathbb{F} \Rightarrow (\alpha f)(v, w) = \alpha f(v, w),$$

其中 $v \in V, w \in W$，称 $L(V, W; \mathbb{F})$ 为**二重线性泛函空间**.

（2）**二重（Z 值）线性映射** 设 V, W, Z 是数域$\mathbb{F}$上的线性空间，若映射 $f: V \times W \to Z$ 对于每个变量都是线性的，亦即

$$\begin{cases} f(\alpha_1 v_1 + \alpha_2 v_2, w) = \alpha_1 f(v_1, w) + \alpha_2 f(v_2, w), \\ f(v, \alpha_1 w_1 + \alpha_2 w_2) = \alpha_1 f(v, w_1) + \alpha_2 f(v, w_2), \end{cases} \quad \alpha_1, \alpha_2 \in \mathbb{F},$$

其中 $v, v_1, v_2 \in V, w, w_1, w_2 \in W$，则称 f 为积空间 $V \times W$ 上的**二重（Z 值）线性映射**，记这样映射（注意上式为 Z 中元素的等式）全体为 $L(V, W; Z) \equiv L(V \times W, Z)$，它在以下运算下构成$\mathbb{F}$上的线性空间：

$$f, g \in L(V, W; Z) \Rightarrow (f+g)(v, w) = f(v, w) + g(v, w);$$
$$f \in L(V, W; Z), \alpha \in \mathbb{F} \Rightarrow (\alpha f)(v, w) = \alpha f(v, w),$$

其中，$v \in V, w \in W$，称 $L(V, W; Z)$ 为**二重线性映射空间**.

（3）r **重（$\mathbb{F}$ 值）线性泛函** 设 $V_1, V_2, \cdots, V_r$ 是$\mathbb{F}$上的线性空间，若 $f: V_1 \times V_2 \times \cdots \times V_r \to \mathbb{F}$ 满足：$f(v_1, v_2, \cdots, v_r)$ 对于每个变量，$v_j \in V_j (j=1,2,\cdots,r)$ 都是线性的，则称 f 为积空间 $V = V_1 \times V_2 \times \cdots \times V_r$ 上的 r **重（$\mathbb{F}$ 值）线性泛函**，记这样的泛函的全体为 $L(V_1, V_2, \cdots, V_r; \mathbb{F}) \equiv L(V_1 \times V_2 \times \cdots \times V_r, \mathbb{F})$，它在以下运算下构成$\mathbb{F}$上的线性空间：

$$f, g \in L(V_1, V_2, \cdots, V_r; \mathbb{F}) \Rightarrow (f+g)(v_1, v_2, \cdots, v_r) = f(v_1, v_2, \cdots, v_r) + g(v_1, v_2, \cdots, v_r);$$
$$f \in L(V_1, V_2, \cdots, V_r; \mathbb{F}), \alpha \in \mathbb{F} \Rightarrow (\alpha f)(v_1, v_2, \cdots, v_r) = \alpha f(v_1, v_2, \cdots, v_r);$$

称 $L(V_1, V_2, \cdots, V_r; \mathbb{F})$ 为 r **重线性泛函空间**（r-**multi-linear functional space**）.

（4）r **重（Z 值）线性映射** 设 $V_1, V_2, \cdots, V_r, Z$ 是$\mathbb{F}$上的线性空间，若映射 $f: V_1 \times \cdots \times V_r \to Z$ 满足：$f(v_1, v_2, \cdots, v_r)$ 对于每个变量 $v_j \in V_j (j=1,2,\cdots,r)$ 都是线性的，则称 f 为积空间 $V_1 \times V_2 \times \cdots \times V_r$ 上的 r **重（Z 值）线性映射**，记这样映射全体为 $L(V_1, V_2, \cdots, V_r; Z) \equiv$

$L(V_1\times V_2\times\cdots\times V_r,Z)$，它在以下运算下构成$\mathbb{F}$上的线性空间：

$$f,g\in L(V_1,V_2,\cdots,V_r;Z)\Rightarrow(f+g)(v_1,v_2,\cdots,v_r)=f(v_1,v_2,\cdots,v_r)+g(v_1,v_2,\cdots,v_r);$$

$$f\in L(V_1,V_2,\cdots,V_r;Z),\alpha\in\mathbb{F}\Rightarrow(\alpha f)(v_1,v_2,\cdots,v_r)=\alpha f(v_1,v_2,\cdots,v_r),$$

称$L(V_1,V_2,\cdots,V_r;Z)$为 r **重线性映射空间**（r-**multi-linear mapping space**）.

3. 多重线性映射空间的同构

对于多重线性空间$L(V_1,V_2,\cdots,V_r;Z)$，这里给出它的同构空间.

1）线性映射空间 $L(V,Z)$ 的同构性

设V,Z分别为数域$\mathbb{F}$上的n维与m维线性空间，基分别为$\mathfrak{A}=\{a_1,a_2,\cdots,a_n\}$，$\mathfrak{B}=\{b_1,b_2,\cdots,b_m\}$. 考虑$f\in L(V,Z)$，$x\in V\to y=f(x)\in Z$，设$f(x)=y$将两基的关系表示为

$$b_j=\sum_{k=1}^{n}f_{jk}a_k,\quad j=1,2,\cdots,m,\tag{2.2.7}$$

记$A=A_{m\times n}=[f_{jk}]_{1\leqslant j\leqslant m,1\leqslant k\leqslant n}$，并用已经熟悉的表示法，由$\mathfrak{A}=\begin{bmatrix}a_1\\a_2\\\vdots\\a_n\end{bmatrix}\equiv M_{\mathfrak{A}}$，$\mathfrak{B}=\begin{bmatrix}b_1\\b_2\\\vdots\\b_m\end{bmatrix}\equiv M_{\mathfrak{B}}$，

$A=M_{\mathfrak{AB}}$，则(2.2.7)式表示为$M_{\mathfrak{B}}\equiv M_{\mathfrak{AB}}M_{\mathfrak{A}}$，并且由定理2.2.1，$L(V,Z)$与数域$\mathbb{F}$上的$m\times n$矩阵空间$\mathfrak{M}_{m\times n}$同构，我们得到$L(V,Z)\xleftrightarrow{\text{同构}}\mathfrak{M}_{m\times n}$.

2）二重线性映射空间 $L(V,W;Z)$ 的分解

希望把二重线性映射空间$L(V,W;Z)$与一个线性空间$L(Y,Z)$对应起来，亦即，要构造一个线性空间Y，使得$f=g\circ h:V\times W\to Z$.

设V,W为$\mathbb{F}$上的n维与m维线性空间，基为$\mathfrak{A}=\{a_1,a_2,\cdots,a_n\}$，$\mathfrak{B}=\{b_1,b_2,\cdots,b_m\}$，$Z$也是$\mathbb{F}$上的线性空间. 我们的目的是把二重线性映射$f\in L(V,W;Z)$转换成双线性映射$h\in L(V\times W,Y)$与线性映射$g\in L(Y,Z)$的复合，这个$Y$与$V$和$W$有关，它满足交换图2.2.3.

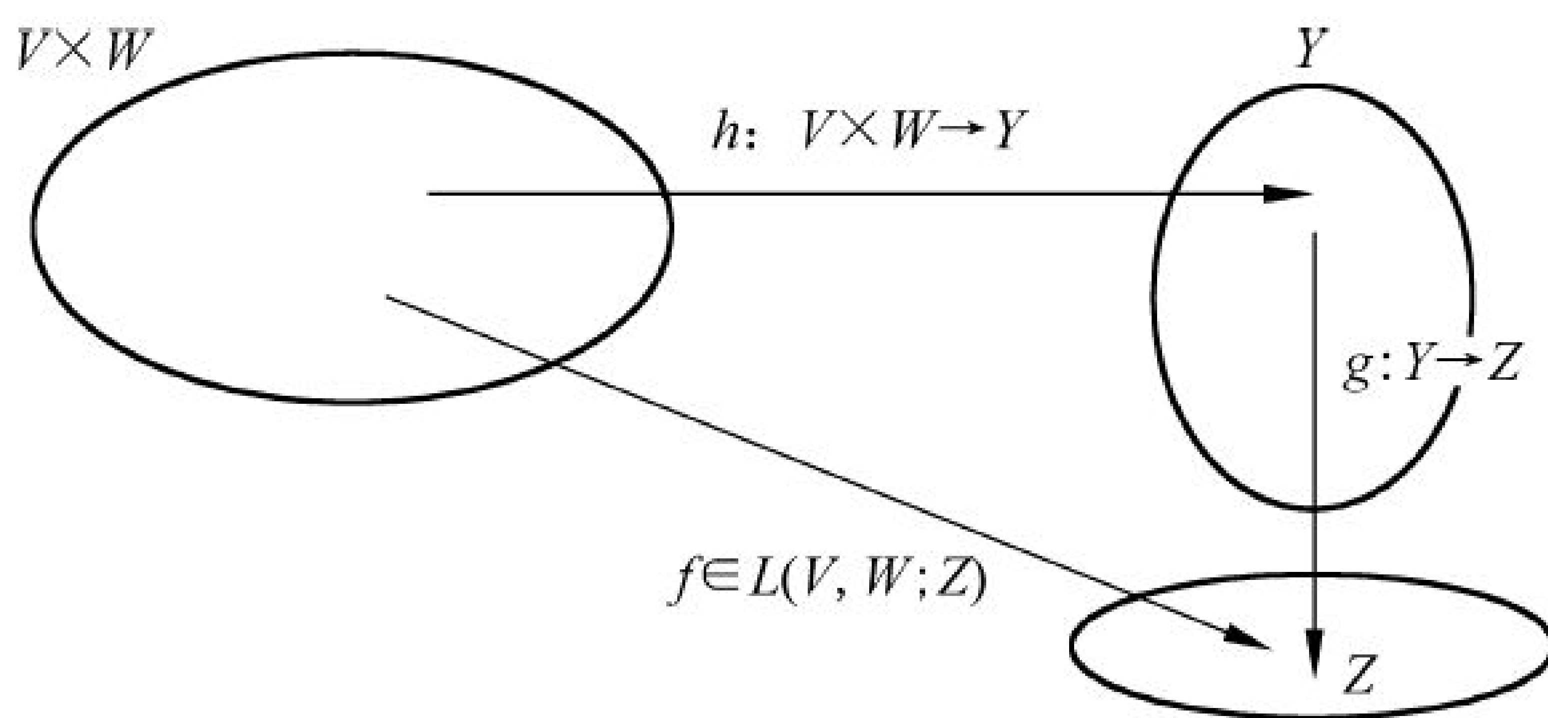


图2.2.3 交换图

下面构造 Y,使得交换图 2.2.3 成立.

4. 张量积与张量积空间

1) V^* 与 W^* 中的张量积

定义 2.2.9(**元 $v^*\in V^*$ 与 $w^*\in W^*$ 的张量积、空间 V^* 与 W^* 的张量积**) 设线性空间 V,W 的对偶空间分别为 V^*,W^*,取 $v^*\in V^*=L(V,\mathbb{F}),w^*\in W^*=L(W,\mathbb{F})$.

(1) 元 $v^*\in V^*$ 与 $w^*\in W^*$ 的**张量积**(**tenser product**) $v^*\otimes w^*:V\times W\to\mathbb{F}$ 满足

$$v^*\otimes w^*(v,w)=v^*(v)w^*(w)=\langle v,v^*\rangle\langle w,w^*\rangle,\quad v\in V,\ w\in W,\qquad(2.2.8)$$

其中 $\langle v,v^*\rangle$ 由定义 2.2.7 给出. 易见,张量积 $v^*\otimes w^*$ 是 $V\times W$ 上的双线性泛函,即 $v^*\otimes w^*\in L(V,W;\mathbb{F})$. 于是,运算$\otimes$视为积空间 $V^*\times W^*$ 到 $L(V,W;\mathbb{F})$上的双线性映射,即

$$\otimes:(v^*,w^*)\in(V^*,W^*)\equiv V^*\times W^*\to v^*\otimes w^*\in L(V,W;\mathbb{F}).$$

(2) 空间 V^* 与空间 W^* 的张量积定义为由形如 $v^*\otimes w^*$ 的元 $v^*\in V^*$ 与元 $w^*\in W^*$ 的张量积所张成的线性空间,亦即

$$V^*\otimes W^*\equiv\operatorname{span}\{v^*\otimes w^*:v\in V^*,w^*\in W^*\},\qquad(2.2.9)$$

此中元的运算定义为 $\forall f,g\in V^*\otimes W^*,\alpha\in\mathbb{F}$,有

$$(f+g)((v,w))=f((v,w))+g((v,w)),\quad(v,w)\in V\times W;$$
$$(\alpha\cdot f)((v,w))=\alpha f((v,w)),\quad(v,w)\in V\times W.$$

注 $V^*\otimes W^*$ 中的元都是由形如 $v^*\otimes w^*$ 的有限线性组合所得到,其表示式相当复杂,$v^*\otimes w^*$ 仅是空间中的“单项式”.

2) V^* 与 W^* 的张量积的表示

设 $\mathfrak{A}=\{a_k,1\leqslant k\leqslant n\}$是 n 维线性空间 V 的基,$\mathfrak{A}^*=\{a_j^*,1\leqslant j\leqslant n\}$是 V^* 的关于基 $\mathfrak{A}$的对偶基,满足 $a_j^*(a_k)=\delta_{kj}(1\leqslant j,k\leqslant n)$. 相应地,设 m 维线性空间 W 的基为 $\mathfrak{B}=\{b_k,1\leqslant k\leqslant m\}$,且 $\mathfrak{B}^*=\{b_j^*,1\leqslant j\leqslant m\}$是 W^* 的关于基 $\mathfrak{B}$的对偶基,满足 $b_j^*(b_k)=\delta_{kj}(1\leqslant j,k\leqslant m)$.

对于运算$\otimes:(v^*,w^*)\in V^*\times W^*\to v^*\otimes w^*\in L(V,W;\mathbb{F})$,由于其双线性性,有

$$v^*\otimes w^*=\sum_{j,k}(v^*(a_j)w^*(b_k))(a_j^*\otimes b_k^*),$$

于是,可以证明$\{a_j^*\otimes b_k^*:1\leqslant j\leqslant n,1\leqslant k\leqslant m\}$是 $V^*\otimes W^*$ 的基,而 $V^*\otimes W^*$ 是$\mathbb{F}$ 上的 $n\times m$ 维线性空间,并且 $V^*\otimes W^*\xleftrightarrow{\text{同构}}L(V,W;\mathbb{F})$,这个对应关系由映射$(v^*,w^*)\in V^*\times W^*\to v^*\otimes w^*\in L(V,W;\mathbb{F})$ 给出.

3) V 与 W 中的张量积

为定义 V 与 W 的张量积,只需考虑 $V=(V^*)^*$ 与 $W=(W^*)^*$,按照 1)与 2)中的方法便可定义 $v\in V,w\in W$ 的张量积 $v\otimes w$;进而定义张量积空间 $V\otimes W$,并且有

$$V\otimes W\xleftrightarrow{\text{同构}}L(V^*,W^*;\mathbb{F}).$$

可以证明，$\{a_j\otimes b_k:1\leqslant j\leqslant n,1\leqslant k\leqslant m\}$是 $V\otimes W$ 的基，而 $V\otimes W$ 是$\mathbb{F}$上的 $n\times m$ 维线性空间.

4）张量积空间的对偶性

令$\langle v\otimes w,v^*\otimes w^*\rangle=\langle v,v^*\rangle\langle w,w^*\rangle$，并且

$$\langle a_j\otimes b_k,a_l^*\otimes b_s^*\rangle=\delta_{jl}\delta_{ks}=\begin{cases}1, & (j,k)=(l,s),\\ 0, & (j,k)\neq(l,s),\end{cases} \tag{2.2.10}$$

从而$\{a_j\otimes b_k:1\leqslant j\leqslant n,1\leqslant k\leqslant m\}$是 $V\otimes W$ 的基，$\{a_j^*\otimes b_k^*:1\leqslant j\leqslant n,1\leqslant k\leqslant m\}$是 $V^*\otimes W^*$ 的基，并且互为对偶基，于是 $V^*\otimes W^*=(V\otimes W)^*$.

将它们的基、对偶基列为表 2.2.1.

表 2.2.1

空间	基	对偶空间	对偶基
V	$\mathfrak{A}=\{a_1,a_2,\cdots,a_n\}$，$\dim V=n$	V^*	$\mathfrak{A}=\{a_1^*,a_2^*,\cdots,a_n^*\}$，$\dim V^*=n$
W	$\mathfrak{B}=\{b_1,b_2,\cdots,b_m\}$，$\dim V=m$	W^*	$\mathfrak{B}=\{b_1^*,b_2^*,\cdots,b_m^*\}$，$\dim W^*=m$
$V\otimes W$	$\{a_j\otimes b_k:1\leqslant j\leqslant n,1\leqslant k\leqslant m\}$ $\dim(V\otimes W)=n\times m$	$V^*\otimes W^*$	$\{a_j^*\otimes b_k^*:1\leqslant j\leqslant n,1\leqslant k\leqslant m\}$ $\dim(V^*\otimes W^*)=n\times m$

定理 2.2.15 设 $h:V\times W\to V\otimes W$ 是张量积$\otimes$定义的双线性映射，即

$$h(v,w)=v\otimes w:V\times W\to V\otimes W,$$

则对于任意的双线性映射 $f:V\times W\to Z$，存在惟一的线性映射 $g:V\otimes W\to Z$，使得

$$f=g\circ h:V\times W\to Z.$$

证 对于任意双线性映射 $f:V\times W\to Z$，构造一个线性映射 $g:V\otimes W\to Z$，使得它在 $V\otimes W$ 的基$\{a_j\otimes b_k:1\leqslant j\leqslant n,1\leqslant k\leqslant m\}$之下的表示为

$$g(a_j\otimes b_k)=f(a_j,b_k),\quad 1\leqslant j\leqslant n,1\leqslant k\leqslant m.$$

于是，对于任意 $v=\sum_{j=1}^{n}v_ja_j\in V$ 与 $w=\sum_{k=1}^{m}w_kb_k\in W$，有

$$g(v\otimes w)=\sum_{j,k}v_jw_kg(a_j\otimes b_k)=\sum_{j,k}v_jw_kf(a_j,b_k)=f(v,w),$$

故 $g:V\otimes W\to Z$ 惟一确定；g 的线性性是明显的.

定理 2.2.16 线性空间 $L(V,W;Z)$与 $L(V\otimes W,Z)$同构.

证 定义映射 $\varphi:L(V\otimes W,Z)\to L(V,W;Z)$，满足

$$\varphi(g)=g\circ h,\quad g\in L(V\otimes W,Z),$$

其中 $h(v,w)=v\otimes w:V\times W\to V\otimes W$，且 $\varphi(g)=g\circ h\in L(V,W;Z)$，因为

$$\varphi(g)=g\circ h:(v,w)\in V\times W\xrightarrow{h}h(v,w)\in V\otimes W\xrightarrow{g}Z.$$

下面证明 $\varphi: L(V\otimes W,Z)\to L(V,W;Z)$ 是一个同构映射. 对于 $(v,w)\in V\times W$,先取定 $h:(v,w)\to v\otimes w$,于是有

(1) $\varphi: g\in L(V\otimes W,Z)\to g\circ h\in L(V,W;Z)$　图 2.2.4 给出了映射

$$\varphi: g\in L(V\otimes W,Z)\to g\circ h\in L(V,W;Z);$$

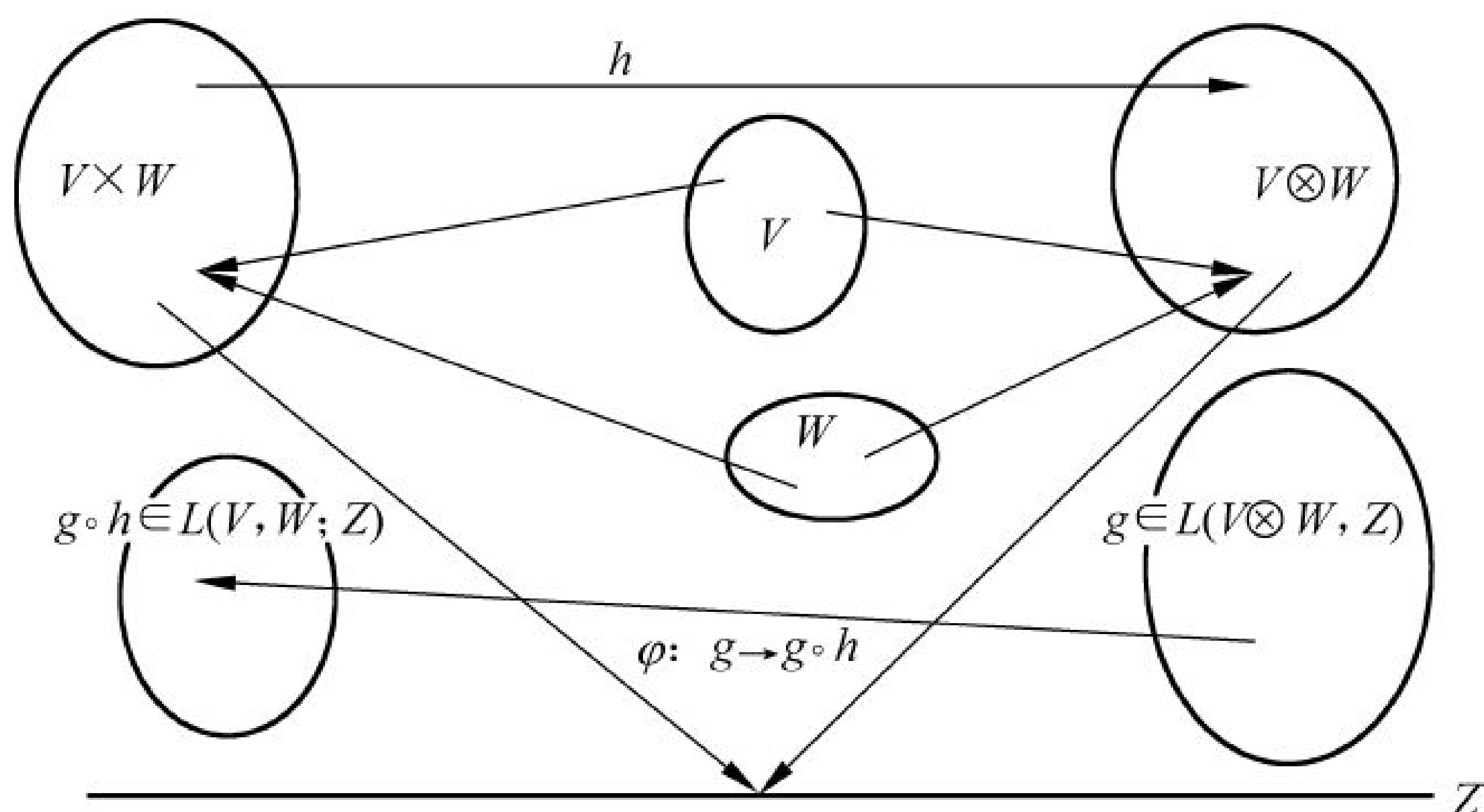


图 2.2.4　同构图

(2) $\varphi: g\to g\circ h$ 是一对一映射　$g_1\neq g_2\in L(V\otimes W,Z)\Rightarrow g_1\circ h\neq g_2\circ h\in L(V,W;Z)$;

(3) $\varphi: g\to g\circ h$ 保持运算

① $g_1,g_2\in L(V\otimes W,Z)\Rightarrow \varphi(g_1+g_2)=(g_1+g_2)\circ h=g_1\circ h+g_2\circ h=\varphi(g_1)+\varphi(g_2)$,这是因为 $\forall(v,w)\in V\times W$,有 $\varphi(g_1+g_2)(v,w)=((g_1+g_2)\circ h)(v,w)=(g_1+g_2)(h(v,w))=g_1(h(v,w))+g_2(h(v,w))=(g_1\circ h)(v,w)+(g_2\circ h)(v,w)=(\varphi(g_1)+\varphi(g_2))(v,w)$.

② $g\in L(V\otimes W,Z),\alpha\in\mathbb{F}\Rightarrow\varphi(\alpha g)=(\alpha g)\circ h=\alpha(g\circ h)=\alpha\varphi(g)$,这是因为 $\forall(v,w)\in V\times W$,有 $\varphi(\alpha g)(v,w)=((\alpha g)\circ h)(v,w)=(\alpha g)(h(v,w))=\alpha g(h(v,w))=\alpha(g\circ h)(v,w)=\alpha\varphi(g)(v,w)$. 定理得证.

5) 张量积的性质

定义 2.2.10(多重线性泛函的张量积)　对于多重线性泛函 $f\in L(V_1,V_2,\cdots V_r;\mathbb{F})$,$g\in L(W_1,W_2,\cdots,W_s;\mathbb{F})$,定义它们的张量积为

$$(f\otimes g)(v_1,v_2,\cdots,v_r,w_1,w_2,\cdots,w_s)=f(v_1,v_2,\cdots,v_r)g(w_1,w_2,\cdots,w_s),$$

其中 $(v_1,v_2,\cdots,v_r)\in V_1\times V_2\times\cdots\times V_r$,$(w_1,w_2,\cdots,w_s)\in W_1\times W_2\times\cdots\times W_s$. 于是,张量积 $f\otimes g$ 是线性空间 $V_1\times\cdots\times V_r\times W_1\times\cdots\times W_s$ 上的 $\mathbb{F}$ 值 $(r+s)$ 重线性泛函,而张量积运算 $\otimes$ 则是线性空间 $L(V_1,\cdots,V_r;\mathbb{F})\times L(W_1,\cdots,W_s;\mathbb{F})$ 到 $L(V_1,\cdots,V_r,W_1,\cdots,W_s;\mathbb{F})$ 的双线性映射.

进而,对于 F 上的线性空间 $V_1,V_2,\cdots,V_r$,同样可以定义张量积空间 $V_1\otimes V_2\otimes\cdots\otimes V_r$,且

$$V_1\otimes V_2\otimes\cdots\otimes V_r\xleftrightarrow{\text{同构}}L(V_1^*,V_2^*,\cdots,V_r^*;\mathbb{F}),$$

$$V_1^* \otimes V_2^* \otimes \cdots \otimes V_r^* \xleftrightarrow{\text{同构}} L(V_1, V_2, \cdots, V_r; \mathbb{F}).$$

定理 2.2.17 张量积运算$\otimes$满足结合律. 对于 $f \in L(V_1, V_2, \cdots, V_r; \mathbb{F})$, $g \in L(W_1, W_2, \cdots, W_s; \mathbb{F})$, $h \in L(U_1, U_2, \cdots, U_t; \mathbb{F})$, 有

$$(f \otimes g) \otimes h = f \otimes (g \otimes h);$$

进而, 设 $h: V_1 \times V_2 \cdots \times V_r \to V_1 \otimes V_2 \otimes \cdots \otimes V_r$ 是由张量积定义的 r 重线性映射, 亦即

$$h(v_1, v_2, \cdots, v_r) = v_1 \otimes v_2 \otimes \cdots \otimes v_r,$$

则对于任意 $f \in L(V_1, V_2, \cdots, V_r; \mathbb{F})$, 存在惟一的线性映射 $g \in L(V_1 \otimes V_2 \otimes \cdots \otimes V, \mathbb{F})$, 使得

$$f = g \circ h: V_1 \times V_2 \times \cdots \times V_r \to \mathbb{F}.$$

习题 2

1. 检验下列集合对于所指的线性运算是否构成实数域上的线性空间:

(1) 次数等于 $n(n \geqslant 1)$ 的实系数多项式全体, 对于多项式的加法和数乘;

(2) 全体 n 阶实对称(反对称, 上三角)方阵, 对于矩阵的加法和数乘;

(3) 平面上的全体向量, 对于通常的加法与如下定义的数乘: $k \cdot \boldsymbol{\alpha} = \mathbf{0}$;

(4) 全体正实数$\mathbb{R}^+$, 加法与数乘定义如下: $a \oplus b = ab, k \cdot a = a^k$.

2. 试证: 线性空间的维数的定义 2.1.2 是合理的, 并求上题中(4)所定义的线性空间的维数以及它的一组基.

3. 设 $A \in \mathfrak{M}_n$, 试证明全体与 A 可交换的方阵构成 $\mathfrak{M}_n$ 的一子空间, 记为 $C(A)$; 且当 $A = I$(单位方阵)时, 求 $C(A)$ 及其维数.

4. 试证: 和 $V_1 + V_2$ 是直和的充分必要条件是 $V_1 \cap V_2 = \{0\}$.

5. 当域$\mathbb{F} = \mathbb{R}$ 时, 证明定理 2.1.2.

6. 设 σ 是数域$\mathbb{F}$ 上的 n 维线性空间$(X, +, \alpha \cdot)$到自身的一个线性映射, 试证:

(1) 对于 X 上的线性泛函 $f \in X^*$, 则 $f \circ \sigma$ 仍然是 X 上的线性泛函;

(2) 定义 X^* 到自身的映射 $\sigma^*: V^* \to V^*$, 满足 $\sigma^*: f \to f \circ \sigma$, 则 σ^* 是 V^* 上的线性映射.

7. 对于 X, X^*, X^{**}, 试证定理 2.1.7 中定义的 $\mathfrak{B} = \{e_1^*, e_2^*, \cdots, e_n^*\}$, $\mathfrak{B}^* = \{e_1^{**}, e_2^{**}, \cdots, e_n^{**}\}$ 分别是 X^*, X^{**} 的基. 又若 $x \to x^{**}$ 是自然同构, 试证: $x^{**}(w) = w(x), \forall w \in X^*$.

8. 试证: 线性空间$(X, +, \alpha \cdot)$上的双线性泛函 $f(x, y)$为斜对称的, 当且仅当 $f(x, x) = 0, \forall x \in X$.

9. 设 V 是 n 维欧氏空间$\mathbb{R}^n$, 其内积为(x, y). 对于 V 中一个确定的向量 a, 定义 V 上的一个泛函 a^*: $\forall b \in V \Rightarrow a^*(b) = (a, b)$. 试证: V 到 V^* 的映射 $a \to a^*$ 是一个同构映射, 在此同构之下, 欧氏空间$\mathbb{R}^n$ 成为自身的对偶空间.

10. 设 σ 是线性空间 V 到自身的线性映射, 若 $\sigma^{k-1}(\xi) \neq 0$, 但 $\sigma^k(\xi) = 0$. 求证: 对于 $k > 1$, 有 $\xi, \sigma(\xi), \sigma^2(\xi), \cdots, \sigma^{k-1}(\xi)$线性无关.

11. 设 V 是复数域$\mathbb{C}$ 上的线性空间, 其维数为 $n \geqslant 2$, 且设 $f(x, y)$是 V 上的对称双线性泛函. 试证:

(1) V 中有非零向量 $\xi \in V$, 使得 $f(\xi, \xi) = 0$;

(2) 若 $f(x,y)$是非退化的,则必有线性无关的向量 ξ,η,满足 $f(\xi,\eta)=1$,$f(\xi,\xi)=f(\eta,\eta)=0$.

12. 考虑线性空间是有限维的内积空间时,共轭算子的两种定义的关系与性质.

13. 验证张量积运算$\otimes$为积空间 $V^*\times W^*$ 到 $L(V,W;F)$上的双线性映射.

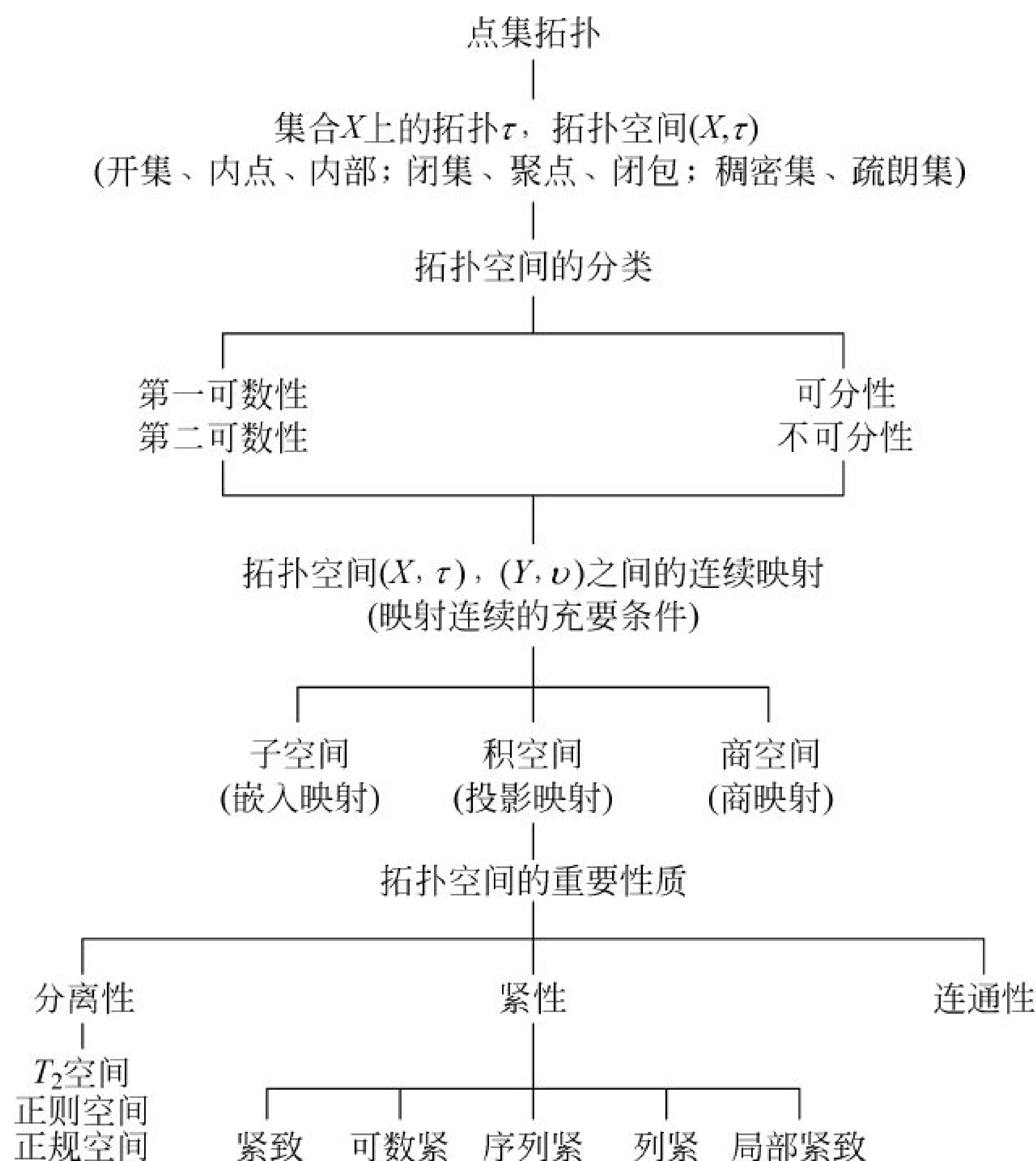
14. 设 V 是数域$\mathbb{F}$上的一个线性空间,$f_1,f_2,\cdots,f_k$ 是 V 上的 k 个线性泛函,试证:

(1) 集合 $W=\{v\in V:f_j(v)=0,1\leqslant j\leqslant k\}$是 V 的一个子空间,称 W 为泛函 $f_1,f_2,\cdots,f_k$ 的**零化子空间**;

(2) V 的任一个子空间皆为某些线性泛函的零化子空间.

第3章 点集拓扑的基本知识

关于点集拓扑，先给出一个框图：



刻画集合中的元素之间的位置关系、远近程度，需要赋予集合以拓扑结构，使集合成为拓扑空间. 拓扑空间在数学科学中占有重要地位，是近代数学中不可缺少的概念和工具. 本章主要参考文献为[2]，[10]，[12]，[14]，[17].

三维欧氏空间$\mathbb{R}^3$ 是人类生存的空间，当然也是人们最熟悉的空间. 平面是二维欧氏空间$\mathbb{R}^2$. 直线为数轴，即一维欧氏空间$\mathbb{R}$. 我们以欧氏空间作为模型来推广空间概念.

在第 2 章中看到，$\mathbb{R}$，$\mathbb{R}^2$，$\mathbb{R}^3$ 在向量加法＋与数乘 $\alpha\cdot$ 之下成为数域$\mathbb{R}$ 或$\mathbb{C}$ 上的线性空

间，这是它们的运算结构. 它们还有一种重要的结构，那就是拓扑结构.

对于$\mathbb{R}$中的任意两点 x,y，定义它们的距离为 $\rho(x,y)=|x-y|$，距离表示了两点之间的远近关系. 类似地，$\mathbb{R}^n(n\in\mathbb{Z}^+)$中任意两点 $x=(x_1,x_2,\cdots,x_n)$，$y=(y_1,y_2,\cdots,y_n)$之间的距离 $\rho(x,y)$定义为 $\rho(x,y)=\sqrt{(x_1-y_1)^2+(x_2-y_2)^2+\cdots+(x_n-y_n)^2}$.

还可以举出许多这样的例子，它们不是欧氏空间，但却可以引进距离概念来刻画元之间的远近程度. 例如，对于数列集合 $X=\{x:x=(x_1,x_2,\cdots,x_n\cdots),x_j\in\mathbb{R},j\in\mathbb{Z}^+\}$，它有子集 $l^2=\left\{x\in X:\sum_{j=1}^{+\infty}|x_j|^2<+\infty\right\}\subset X$，对于 $x,y\in l^2$，可以定义距离

$$\rho(x,y)=\sqrt{\sum_{j=1}^{+\infty}(x_j-y_j)^2}.$$

在 $\mathbb{R},\mathbb{R}^2,\mathbb{R}^3$ 中距离有直观意义，但在其他集合中未必如此. 究竟什么是距离，在数学科学中一定要有明确的定义、明确的特征性质，才能用它来刻画集合中两个元素之间的远近关系，也才能把距离推广到一般集合上去.

3.1 度量空间、赋范线性空间

3.1.1 度量空间

欧氏空间$\mathbb{R},\mathbb{R}^2,\mathbb{R}^3$ 中的距离 $\rho(x,y)$满足什么特征性质呢？它满足下面三个特征性质：

(1) 任意两点间的距离是非负的，即 $\rho(x,y)\geqslant 0$. 并且一点到自身的距离为零，亦即 $y=x\Leftrightarrow\rho(x,y)=0$；

(2) 点 A 到点 B 的距离与点 B 到点 A 的距离相等，即 $\rho(x,y)=\rho(y,x)$；

(3) 三角形的两边之和大于第三边，即 $\rho(x,y)<\rho(x,z)+\rho(z,y)$.

从这三个特征性质出发，可以从更抽象的角度来定义空间中的距离.

定义 3.1.1(度量空间) 设 X 是一个集合，其元为 $x,y,\cdots$，如果 X 中的任意两个元 x，y 都对应于一个实数，记为 $\rho(x,y)$，具有如下性质：

(1) $\rho(x,y)\geqslant 0$, $\rho(x,y)=0\Leftrightarrow x=y$； (非负性)

(2) $\rho(x,y)=\rho(y,x)$； (对称性)

(3) $\rho(x,y)\leqslant\rho(x,z)+\rho(z,y)$， (三点不等式)

则称 $\rho(x,y)$为 x 与 y 之间的**距离(distance)**，称 X 为一个**度量空间(metric space)**. 此时，称集合 X 具有由距离 $\rho(x,y)$赋予的**拓扑结构(topological structure)**.

在很多参考书上，也称度量空间为距离空间.

注 “三点不等式”是“三角形两边之和大于第三边”的推广，前者包含了三点可以在一直线上的情况.

除了欧式空间$\mathbb{R},\mathbb{R}^2,\cdots,\mathbb{R}^n$，下面再看几个重要的度量空间的例子.

1）l^2 空间

给定集合

$$l^2=\left\{x=(x_1,x_2,\cdots,x_n\cdots):x_j\in\mathbb{R},j\in\mathbb{Z}^+,\sum_{j=1}^{+\infty}|x_j|^2<+\infty\right\}. \tag{3.1.1}$$

在 l^2 中定义

$$\rho(x,y)=\sqrt{\sum_{j=1}^{+\infty}(x_j-y_j)^2}, \tag{3.1.2}$$

不难验证，$\rho(x,y)=\sqrt{\sum_{j=1}^{+\infty}(x_j-y_j)^2}$ 满足定义 3.1.1 中的三个性质，因此，它是集合 l^2 上的距离，而 l^2 成为一个度量空间. 有时，为了强调它的拓扑结构，采用记号 $(l^2,\rho(x,y))$，在没有歧义的情况下，用 l^2 就可以了.

2）区间$[a,b]$上的连续函数空间 $C([a,b])$

区间$[a,b]$上实值连续函数的全体所成的集合记为

$$C([a,b])=\{f:f(x)\ \text{是}[a,b]\ \text{上的连续函数}\}. \tag{3.1.3}$$

对于 $f,g\in C([a,b])$，定义

$$\rho(f,g)=\max_{x\in[a,b]}|f(x)-g(x)|, \tag{3.1.4}$$

不难验证，$\rho(f,g)$是 $C([a,b])$上的距离，使得 $C([a,b])$成为一个度量空间.

3）L 可积函数空间 $L^1(E)$、L 平方可积函数空间 $L^2(E)$

$$L^1(E)=\left\{f:\int_E|f(x)|\,\mathrm{d}m<+\infty\right\},\quad L^2(E)=\left\{f:\int_E|f(x)|^2\mathrm{d}m<+\infty\right\}. \tag{3.1.5}$$

分别在 $L^1(E)$与 $L^2(E)$中引入

$$\rho(f,g)=\int_E|f(x)-g(x)|\,\mathrm{d}m,\quad \rho(f,g)=\left\{\int_E|f(x)-g(x)|^2\mathrm{d}m\right\}^{\frac{1}{2}}, \tag{3.1.6}$$

不难验证，它们都满足距离的三条基本性质.

下面以验证 $L^2(E)$是度量空间为例. 对于(1)，任取 $f,g\in L^2(E)$，由 Lebesgue 积分的定义，显然有

$$\rho(f,g)=\left\{\int_E|f(x)-g(x)|^2\mathrm{d}m\right\}^{\frac{1}{2}}\geqslant 0;$$

$f=g$ 蕴含 $\rho(f,g)=0$ 也是显然的. 反之，当 $\rho(f,g)=\left\{\int_E|f(x)-g(x)|^2\mathrm{d}m\right\}^{\frac{1}{2}}=0$ 时，用反证法可以得到：对于几乎处处的 $x\in E$，有 $f=g$（“几乎处处相等”是指：f 与 g 只在 E 的一个零测度子集上不相等）. 这里约定，在 $L^2(E)$中，几乎处处(almost everywhere，简记为 a. e.)相等的函数视为同一个函数. 故定义 3.1.1 中的(1)成立. 对于(2)，对称性显然.

至于定义 3.1.1 中的(3)，由 Hölder 不等式，若 $f,g\in L^2(E)$，则

$$\int_E |f(x)g(x)|\,\mathrm{d}m \leqslant \left\{\int_E |f(x)|^2\mathrm{d}m\right\}^{\frac{1}{2}}\left\{\int_E |g(x)|^2\mathrm{d}m\right\}^{\frac{1}{2}}, \tag{3.1.7}$$

可证得 Minkovski 不等式

$$\left\{\int_E |f(x)+g(x)|^2\mathrm{d}m\right\}^{\frac{1}{2}} \leqslant \left\{\int_E |f(x)|^2\mathrm{d}m\right\}^{\frac{1}{2}} + \left\{\int_E |g(x)|^2\mathrm{d}m\right\}^{\frac{1}{2}},\quad f,g\in L^2(E) \tag{3.1.8}$$

从而(3)成立.

上述三例在第 1 章中都已经赋予运算结构而成为线性空间，故它们都是既有运算结构又有拓扑结构的空间，可分别记为$(l^2,+,\alpha\cdot,\rho_{l^2}(x,y))$，$(C([a,b]),+,\alpha\cdot,\rho_{C([a,b])}(x,y))$，$(L^1(E),+,\alpha\cdot,\rho_{L^1(E)}(x,y))$，$(L^2(E),+,\alpha\cdot,\rho_{L^2(E)}(x,y))$，简记为$(l^2,+,\alpha\cdot,\rho)$，等. 然而，度量空间中的元却不一定要有运算结构，例子如下.

4) 对集合 $A=\{a,b,c,d\}$，定义 $\rho(x,y)=\begin{cases}1, & x\neq y,\\ 0, & x=y,\end{cases}$ 则 $A=\{a,b,c,d\}$成为一个度量空间(A,ρ_A)，但是元与元之间却没定义运算.

3.1.2 赋范线性空间

欧氏空间$\mathbb{R},\mathbb{R}^2,\cdots,\mathbb{R}^n$ 不仅作为线性空间有运算结构，而且作为度量空间有拓扑结构. 空间 $l^2,C([a,b]),L^1(E),L^2(E)$也是如此. 这就启发我们将集合的拓扑结构与运算结构结合起来，构成一种新型的空间，在这种空间中，对每个元都可引进一个将欧氏空间中向量长度推广的概念.

定义 3.1.2(赋范线性空间) 若在线性空间$(X,+,\alpha\cdot)$中，定义对应于每个元素 $x\in X$ 的非负实数 $\|x\|_X$，具有如下性质：

(1) $\|x\|_X\geqslant 0$，$\|x\|_X=0\Leftrightarrow x=0$($x=0$ 是 X 的零元)； (非负性)

(2) $\|\alpha x\|_X=|\alpha|\,\|x\|_X$，$\alpha\in\mathbb{F}$； (绝对齐性)

(3) $\|x+y\|_X\leqslant\|x\|_X+\|y\|_X$，$x,y\in X$， (三点不等式)

则称 X 为一个**赋范线性空间(normed linear space)**，简称**赋范空间(normed space)**，$\|x\|_X$称为 x 的**范数(norm)**. 此时，线性空间 X 赋予由范数 $\|x\|_X$决定的拓扑结构，记为$(X,+,\alpha\cdot,\|x\|_X)$，在没有歧义时，记为$(X,\|x\|_X)$，或$(X,\|x\|)$，或最简记为 X.

$\mathbb{R},\mathbb{R}^2,\mathbb{R}^3,C([a,b]),L^1(E),L^2(E)$都是赋范线性空间，其范数分别为

$$\|x\|_{\mathbb{R}^n}=\sqrt{x_1^2+x_2^2+\cdots+x_n^2},\quad x\in\mathbb{R}^n;$$

$$\|x\|_{l^2}=\sqrt{\sum_{j=1}^{+\infty}x_j^2},\quad x\in l^2;$$

$$\|f\|_{C([a,b])} = \max_{x\in[a,b]} |f(x)|, \quad f \in C([a,b]);$$

$$\|f\|_{L^1(E)} = \int_E |f(x)| \mathrm{d}m, \quad f \in L^1(E);$$

$$\|f\|_{L^2(E)} = \left\{\int_E |f(x)|^2 \mathrm{d}m\right\}^{\frac{1}{2}}, \quad f \in L^2(E).$$

每一个赋范线性空间是一个度量空间，因为若对赋范线性空间$(X,\|x\|_X)$的任意两个元$x,y\in X$，定义$\rho(x,y)=\|x-y\|_X$，则(X,ρ)就成为一个度量空间. 但反之却不然，因为度量空间未必是线性空间（未必有运算结构）.

在2.1.4节中定义的（实或复）内积空间$(X,+,\alpha\cdot,(x,y))$中，可以定义范数，使其成为赋范线性空间. 事实上，对于$x\in X$，定义范数$\|x\|_X=\{(x,x)\}^{\frac{1}{2}}$，则不难证明，空间$(X,+,\alpha\cdot,\|x\|_X)$是一个赋范线性空间；但反之却不然，许多赋范线性空间上不能定义内积而成为内积空间，例如$C([a,b])$，$L^1(E)$等，它们只是赋范线性空间，却不是内积空间.

3.2 拓扑空间

3.2.1 拓扑空间中的一些定义

人们从研究集合中元之间的位置关系与远近程度的需要出发，引进集合的拓扑结构，如度量空间、赋范空间、内积空间、欧氏空间等. 然而，自然科学领域中所遇到的空间常常会更为一般. 不断在新的高度、深度、广度意义下研究集合的拓扑结构，就逐步形成了"点集拓扑学".

1. 拓扑结构

定义3.2.1（拓扑） 设X是一个非空集合，$X\neq\varnothing$，τ是X的一些子集所成的集族. 若τ满足

(1) $X,\varnothing\in\tau$；

(2) $G_1,G_2\in\tau \Rightarrow G_1\cap G_2\in\tau$； （有限交性质）

(3) $G_\lambda\in\tau,\lambda\in\Lambda \Rightarrow \bigcup_{\lambda\in\Lambda}G_\lambda\in\tau$，$\Lambda$为指标集， （任意并性质）

则称τ为X上的一个**拓扑（topology）**，或称τ决定了X上的一个**拓扑结构**. τ中的任一集$A\in\tau$称为X的**开集（open set）**，并称(X,τ)为**拓扑空间（topological space）**.

定义3.2.1中的(1)表示"整个集合X与空集$\varnothing$是开集"；(2)"有限交性质"表示"有限个开集的交集是开集"；(3)"任意并性质"表示"任意多个开集的并集是开集".

我们约定，**空集（empty set）**$\varnothing$包含在任意集合中，即$\forall A\subset X \Rightarrow \varnothing\subset A$.

作为例子，度量空间、赋范线性空间、内积空间、欧氏空间都是拓扑空间. 但要强调的是，

拓扑空间中不需要有运算结构，正像度量空间中不需要有运算结构一样.

下面来看在度量空间中的拓扑 τ 是如何确定的.

给定度量空间(X,ρ)，对于任一点 $x\in X$ 与任意实数 $r>0$，称集合

$$B(x,r)=\{y\in X:\rho(x,y)<r\} \tag{3.2.1}$$

为度量空间(X,ρ)中以 x 为球心、r 为半径的开球，简称 $B(x,r)$为球.

度量空间(X,ρ)中的集合 $G\subset X$ 称为**开集**，若对于 G 中的任意点 $x\in G$，必存在开球

$$B(x,r)=\{y\in X:\rho(x,y)<r\},$$

使得 $B(x,r)\subset G$.

图 3.2.1 中细虚线所围的是集合 $G\subset X$，粗虚线所围的是球 $B(x,r)$.

度量空间(X,ρ)中的开集的全体记为

$$\tau=\{G\subset X:G\text{ 是}(X,\rho)\text{中的开集}\},$$

它满足定义 3.2.1 中的(1)～(3)，证明如下.

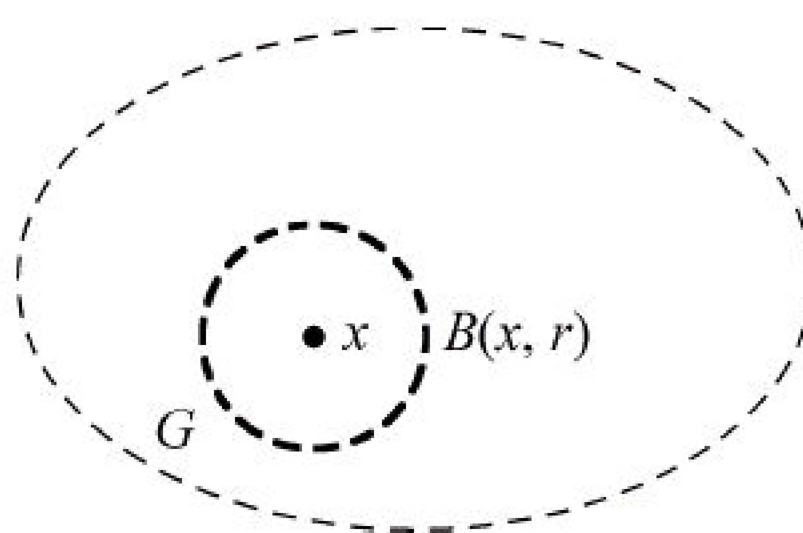


图 3.2.1 度量空间中的开集

(1) 度量空间 X 本身是开集. 因为 $\forall x\in X$，显然 $\exists r>0$(但不惟一)，使得 $B(x,r)=\{y\in X:\rho(x,y)<r\}\subset X$，故 $X\in\tau$；又由约定，$\varnothing\in\tau$.

(2) 设 G_1,G_2 是度量空间(X,ρ)中的开集，则它们满足 $\exists r_1$, s. t. $B(x,r_1)\subset G_1$, $\exists r_2$, s. t. $B(x,r_2)\subset G_2$. 于是有两种情况：① $G_1\cap G_2=\varnothing$，则 $G_1\cap G_2=\varnothing\in\tau$；② $G_1\cap G_2\neq\varnothing$，则取 $r=\min\{r_1,r_2\}$，使得 $\forall x\in G_1\cap G_2\Rightarrow B(x,r)\subset G_1\cap G_2$，故 $G_1\cap G_2\in\tau$.

此性质对于有限多个开集也成立，即 $G_1,G_2,\cdots,G_n\in\tau\Rightarrow\bigcap_{j=1}^{n}G_j\in\tau$.

(3) 设 $G_\lambda,\lambda\in\Lambda$，是度量空间$(X,\rho)$中的开集，$\Lambda$ 是指标集，它可以是有限集，也可以是无限集(可数或不可数). 令 $G=\bigcup_{\lambda\in\Lambda}G_\lambda$. 我们证明 $G=\bigcup_{\lambda\in\Lambda}G_\lambda\in\tau$. 推理如下：

$$\begin{aligned}\forall x\in G&\Rightarrow\exists\lambda_0\in\Lambda,\ \text{s. t. }x\in G_{\lambda_0}\subset\bigcup_{\lambda\in\Lambda}G_\lambda\\&\Rightarrow\text{因 }G_{\lambda_0}\text{ 是开集，故}\exists r_0,\ \text{s. t. }B(x,r_0)\subset G_{\lambda_0}\\&\Rightarrow\exists r_0,\ \text{s. t. }B(x,r_0)\subset G_{\lambda_0}\subset G=\bigcup_{\iota\in\Lambda}G_\iota\\&\Rightarrow G\text{ 是开集.}\end{aligned}$$

于是，τ 成为度量空间(X,ρ)上的拓扑. 所以，一个度量空间(X,ρ)是一个拓扑空间.

作为习题，请读者分别给出赋范线性空间$(X,\|x\|_X)$、内积空间$(X,(x,y))$、欧氏空间 $\mathbb{R}^n$ 中的开集族 τ，使得它们成为拓扑空间.

拓扑空间(X,τ)中，还有一类重要的集合，与开集具有同等重要的地位，那就是闭集.

定义 3.2.2(闭集) 设(X,τ)是拓扑空间，$A\subset X$ 是 X 的子集，若 A 的补集 $A^C\equiv X\backslash A$

是开集，亦即 $A^C\in\tau$，则称 A 为**闭集**(**closed set**). 补集 $A^C\equiv X\backslash A$ 也记为 $\complement A\equiv X\backslash A$.

例如，在 $\mathbb{R}^1\equiv\mathbb{R}$ 中，① 闭区间 $[a,b]$ 是闭集；② 有限多个闭区间的并 $\bigcup_{j=1}^{k}[a_j,b_j]$ $(k\in\mathbb{Z}^+)$ 是闭集.

根据集合的运算规则，可得到闭集满足如下性质.

定理 3.2.1 设 (X,τ) 是拓扑空间，则

(1) $X,\varnothing$ 是闭集；

(2) 若 F_1,F_2 为闭集，则 $F_1\cup F_2$ 为闭集； （有限并性质）

(3) 若 $F_\lambda(\lambda\in\Lambda)$ 为闭集，则 $\bigcap_{\lambda\in\Lambda}F_\lambda$ 为闭集. （任意交性质）

证 对于(1)，因 $X^C=X\backslash X=\varnothing$，由约定 $\varnothing\in\tau$ 是开集，从而 X 是闭集；同理，空集 $\varnothing$ 也是闭集.

对于(2)，由 de Morgan 公式 $X\backslash\left\{\bigcup_{\lambda\in\Lambda}A_\lambda\right\}=\bigcap_{\lambda\in\Lambda}(X\backslash A_\lambda)$，得

F_1,F_2 为闭集 $\Rightarrow X\backslash(F_1\cup F_2)=(X\backslash F_1)\cap(X\backslash F_2)$，其中 $X\backslash F_1,X\backslash F_2$ 为开集 $\Rightarrow$ $X\backslash(F_1\cup F_2)$ 为两开集 $X\backslash F_1,X\backslash F_2$ 的交，故 $X\backslash(F_1\cup F_2)$ 为开集 $\Rightarrow F_1\cup F_2$ 为闭集.

对于(3)，由 de Morgan 公式 $X\backslash\left\{\bigcap_{\lambda\in\Lambda}A_\lambda\right\}=\bigcup_{\lambda\in\Lambda}(X\backslash A_\lambda)$，得

$F_\lambda(\lambda\in\Lambda)$ 为闭集 $\Rightarrow X\backslash\left\{\bigcap_{\lambda\in\Lambda}F_\lambda\right\}=\bigcup_{\lambda\in\Lambda}(X\backslash F_\lambda)$，其中每个 $X\backslash F_\lambda$ 为开集 $\Rightarrow X\backslash\left\{\bigcap_{\lambda\in\Lambda}F_\lambda\right\}$ 为开集族 $X\backslash F_\lambda(\lambda\in\Lambda)$ 的并，故 $X\backslash\left\{\bigcap_{\lambda\in\Lambda}F_\lambda\right\}$ 为开集 $\Rightarrow\bigcap_{\lambda\in\Lambda}F_\lambda$ 为闭集.

2. 点的邻域

定义 3.2.3(邻域) 设 (X,τ) 为拓扑空间，对于点 $x\in X$，称 X 的子集 A 为 x 的**邻域**(**neighborhood**)，若存在开集 $G\in\tau$，使得 $x\in G\subset A$(图 3.2.2).

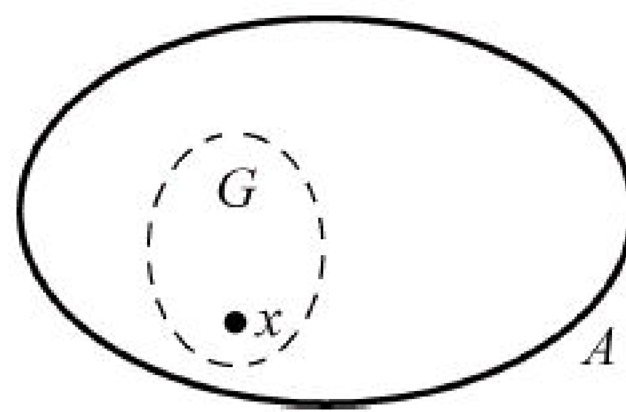


图 3.2.2 拓扑空间中点的邻域

对于拓扑空间 (X,τ)，记点 $x\in X$ 的邻域全体为

$$\mathfrak{U}_x=\{A\subset X: A \text{ 是 } x\in A \text{ 的邻域}\},$$

称 $\mathfrak{U}_x$ 为点 $x\in X$ 的**邻域系**(**neighborhood system**).

易见，包含点 x 的一个开集 $O\in\tau$ 就是 x 的一个邻域，常称为 x 的一个**开邻域**.

定理 3.2.2 设 (X,τ) 是拓扑空间，$\mathfrak{U}_x$ 为点 $x\in X$ 的邻域系，则

(1) $A_1,A_2\in\mathfrak{U}_x\Rightarrow A_1\cap A_2\in\mathfrak{U}_x$； （有限交性质）

(2) $A_\iota\in\mathfrak{U}_x,\iota\in\Lambda\Rightarrow\bigcup_{\iota\in\Lambda}A_\iota\in\mathfrak{U}_x$； （任意并性质）

(3) $\forall A\in\mathfrak{U}_x \Rightarrow \exists B\in\mathfrak{U}_x$, s.t. ① $B\subset A$, ② $\forall y\in B$, 有 $B\in\mathfrak{U}_y$;

(4) $G\in\tau$ 为开集 $\Leftrightarrow \forall x\in G$, 有 $G\in\mathfrak{U}_x$.

证明留作习题.

注 若在集合 X 上定义每个点 $x\in X$ 的邻域系, 则 $\mathfrak{U}=\{\mathfrak{U}_x: x\in X\}$ 同样也能给出集合 X 上的拓扑, 使得 $(X,\mathfrak{U})$ 成为一个拓扑空间.

3. 集合的内点与内部、外点与外部、边界点与边界

1) 集合 $A\subset X$ 的内点 对于点 $x\in A$, 若存在开集 $G\in\tau$, 使得 $x\in G\subset A$, 则称 x 为集合 A 的**内点**(**inner point**).

集合 $A\subset X$ 的内部 集合 A 的内点的全体称为 A 的**内部**(**inner**), 记为 $\mathring{A}$ 或 A°.

2) 集合 $A\subset X$ 的外点 对于点 $x\in A$, 若 $x\in(A^{C})^{\circ}$, 亦即 x 属于 A 的补集的内部, 则称 x 为 A 的**外点**(**external point**).

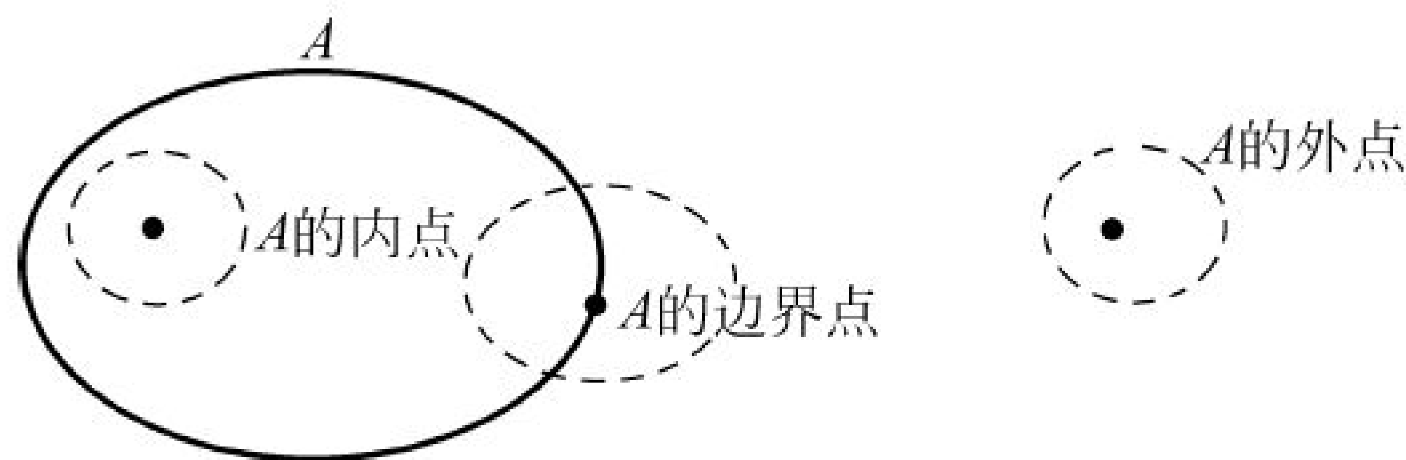


图 3.2.3 拓扑空间中集合 A 的内点、外点、边界点

集合 $A\subset X$ 的外部 集合 A 的外点的全体称为 A 的**外部**(**external**).

3) 集合 $A\subset X$ 的边界点 若在包含点 $a\in X$ 的任意开集 G 中, 既含有 A 中的点、又含有 A 外的点, 则称 a 为 A 的**边界点**(**boundary point**).

集合 $A\subset X$ 的边界 A 的边界点的全体称为 A 的**边界**(**boundary**), 记为 ∂A.

4. 集合的聚点与孤立点、导集与闭包

1) 集合 $A\subset X$ 的聚点 若对包含点 $a\in X$ 的任意开集 $G\in\tau$, 集合 $G\backslash\{a\}$ 与 A 的交集非空, 即 $(G\backslash\{a\})\cap A\neq\varnothing$, 则称 a 为 A 的**聚点**(**accumulation point**), 或称**极限点**(**limit point**);

集合 $A\subset X$ 的孤立点 若 $a\in A$ 但不是 A 的聚点, 则称 a 为 A 的**孤立点**(**isolated point**);

2) 集合 $A\subset X$ 的导集 A 的聚点的全体称为 A 的**导集**(**divided set**), 记为 A';

3) 集合 $A\subset X$ 的闭包 A 与 A 的导集的并集称为的 A 的**闭包**(**closure**), 记为 $\overline{A}$, 即 $\overline{A}=A\cup A'$.

若集合 $A\subset X$ 是闭集, 且没有孤立点, 则称 A 为**完全集**(**complete set**).

注 集合 A 的边界点与聚点都可以属于 A, 也可以不属于 A; 但孤立点一定属于 A.

闭集与闭包有如下性质.

定理 3.2.3 设(X,τ)是拓扑空间，则集合$A\subset X$的闭集A有如下等价性质：

(1) A为闭集 $\Leftrightarrow$ (2)$A'\subset A$ $\Leftrightarrow$ (3)$A=A\cup A'$ $\Leftrightarrow$ (4)$\overline{A}=A$.

证 推理如下：

证(1) $\Leftrightarrow$ (2).

A为闭集 $\Rightarrow$ A^C为开集 $\Rightarrow$ $\forall x\in A^C$，$\exists G\in\tau$且$x\in G\subset A^C$，s. t. $G\cap A=\varnothing$ $\Rightarrow$ $(G\setminus\{x\})\cap A=\varnothing$(因$\forall x\notin A$) $\Rightarrow$ $\forall x\in A^C$不是A的聚点，即$x\notin A'$ $\Rightarrow$ $A'\subset A$(因“$\forall y\in A^C\Rightarrow y\notin A'$”可推得“$\forall y\in A'\Rightarrow y\notin A^C\Rightarrow y\in A$”) $\Rightarrow$ (1)蕴含(2)；

反之，$A'\subset A\Rightarrow(A')^C\supset A^C\Rightarrow\forall x\in A^C$有$x\in(A')^C$，故$x\notin A'$且$x\notin A\Rightarrow\exists G_x\in\tau$，s. t. $(G_x\setminus\{x\})\cap A=\varnothing\Rightarrow(G_x\setminus\{x\})\cap A=G_x\cap A=\varnothing$ (因$x\notin A$) $\Rightarrow A^C=\bigcup\limits_{x\in A^C}\{x\}\subset\bigcup\limits_{x\in A^C}G_x\subset A^C\Rightarrow A^C=\bigcup\limits_{x\in A^C}G_x$为开集的并集 $\Rightarrow$ A为闭集 $\Rightarrow$ (2)蕴含(1).

证(2) $\Leftrightarrow$ (3).

$A'\subset A\Rightarrow A\cup A'\subset A\cup A=A\Rightarrow$ 由$A\subset A\cup A'$即得$A=A\cup A'$ $\Rightarrow$ (2)蕴含(3)；

反之，$A=A\cup A'\Rightarrow A'\subset A$(否则，$\exists x\in A'$，但$x\notin A$，与$A'\subset A$矛盾) $\Rightarrow$ (3)蕴含(2).

证(3) $\Leftrightarrow$ (4).

$\overline{A}=A\Leftrightarrow A\cup A'=A$(定义).

例 3.2.1 (1) 在一维欧氏空间$\mathbb{R}$中，集$A=(-a,a]$的闭包是$\overline{A}=[-a,a]$，内部是$\mathring{A}=(-a,a)$，边界是$\partial A=\{-a,a\}$，外部是$(-\infty,a]\cup(a,+\infty)$，导集是$A'=[-a,a]$.

(2) 若$A_n=\left(\dfrac{1}{n},2\right)$，则$\overline{A_n}=\left[\dfrac{1}{n},2\right]$，$\bigcup\limits_{n=1}^{+\infty}\overline{A_n}=\bigcup\limits_{n=1}^{+\infty}\left[\dfrac{1}{n},2\right]=(0,2]$，

$$\overline{\bigcup_{n=1}^{+\infty}A_n}=\overline{\bigcup_{n=1}^{+\infty}\left(\frac{1}{n},2\right)}=\overline{(0,2)}=[0,2].$$

5. 稠密集

定义 3.2.4(稠密集) 设(X,τ)是拓扑空间，$A\subset X$是X的子集.

(1) 若A的闭包是X，亦即$X=\overline{A}$，则称A为X的**稠密子集**，或称集合A在X中**稠密**(**dense**)；对X的子集$B\subset X$，若$B\subseteq\overline{A}$，则称A**在**B**中稠密**(此时可以有$A\subset B$，也可以没有).

(2) 若A的闭包的内部是空集$\varnothing$，亦即$(\overline{A})^\circ=\varnothing$，则称$A$是$X$的**无处稠密子集**(或称**疏朗集**).

例 3.2.2 有理数集$\mathbb{Q}$在实数集$\mathbb{R}$中是稠密的，$\overline{\mathbb{Q}}=\mathbb{R}$；整数集$\mathbb{Z}$在$\mathbb{R}$中是无处稠密的，因$\mathbb{Z}$的闭包不含任何有限实数，故$\overline{\mathbb{Z}}$的内部是空集.

定理 3.2.4 设(X,τ)是拓扑空间，若子集$A\subset X$是稠密子集，$X=\overline{A}$，则

$$\forall x\in X\Rightarrow\text{任何开集}\ U\in\tau\ \text{且}\ x\in U\text{，必有}\ U\cap A\neq\varnothing.$$

对定理3.2.4的理解是：任意$x\in X$的任意开邻域$v\in\tau$与A的交非空. 取$A=\mathbb{Q}$，$X=$

$\mathbb{R}$ 为例，对任意实数 $x\in\mathbb{R}$ 及包含此 x 的任意开集，例如取 $v=\left(x-\frac{1}{n},x+\frac{1}{n}\right)(n=1,2,\cdots)$，必有

$$\left(x-\frac{1}{n},x+\frac{1}{n}\right)\cap\mathbb{Q}\neq\varnothing,\quad n=1,2,\cdots,$$

这也说明存在有理数点序列 $q_n\in\left(x-\frac{1}{n},x+\frac{1}{n}\right)\cap\mathbb{Q}\neq\varnothing(n=1,2,\cdots)$，使得 $\lim\limits_{n\to+\infty}q_n=x$.

由此可见，对于拓扑空间 (X,τ)，为研究其某些性质，可转化到考虑其稠密子集上去.

3.2.2 拓扑空间的初步分类

对于拓扑空间 (X,τ)，需要进行分类研究. 本小节仅从 (X,τ) 的拓扑基与 (X,τ) 包含可数子集与否来进行初步分类. 今后，还要从各种不同角度对拓扑空间进行分类，以揭示其内在性质.

1. 拓扑基

1) 拓扑空间 (X,τ) 的拓扑基

定义 3.2.5（拓扑空间 (X,τ) 的拓扑基） 设 (X,τ) 是拓扑空间，若 $\mathfrak{B}$ 是 τ 的一个开子集族，且对于 τ 中的每个开集 $G\in\tau$ 以及 $\forall x\in G$，存在 $B\in\mathfrak{B}$ 使得 $x\in B\subset G$（如图 3.2.4 所示），则称 $\mathfrak{B}$ 为拓扑 τ 的**拓扑基（topological basis, basis of topology）**.

拓扑基有如下特征性质.

定理 3.2.5 设 (X,τ) 是拓扑空间，则 $\mathfrak{B}\subset\tau$ 是 (X,τ) 的拓扑基，当且仅当每个开集 $G\in\tau$ 是 $\mathfrak{B}$ 中一些开集的并. 亦即 $\forall G\in\tau$，存在开子集族 $\mathfrak{B}'\subset\mathfrak{B}$ 使得

$$G=\bigcup_{B\in\mathfrak{B}'}B\quad(\text{或 } G=\bigcup\{B:B\in\mathfrak{B}'\}).$$

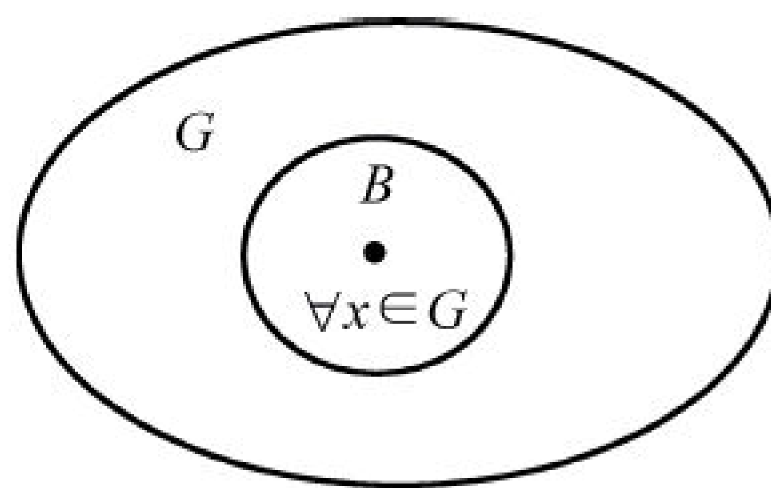


图 3.2.4 拓扑基

证 "⇒" 设 $\mathfrak{B}\subset\tau$ 是 (X,τ) 的拓扑基，则 $\forall G\in\tau, x\in G$，$\exists B\in\mathfrak{B}$ 使得 $x\in B\subset G$. 于是 $G=\bigcup\limits_{x\in B}\{x\}\subset\bigcup\limits_{x\in B}B$. 然而，因每个 B 都含在 G 中，故 $G=\bigcup\limits_{x\in B}\{x\}\subset\bigcup\limits_{x\in B}B\subset G$，从而 $G=\bigcup\limits_{x\in B}B$，这就说明 $\forall G\in\tau$ 可表示为 $\mathfrak{B}$ 中一些开集的并.

"⇐" 对于拓扑空间 (X,τ)，若开子集族 $\mathfrak{B}\subset\tau$ 满足 $\forall O\in\tau$，$\exists\mathfrak{B}'\subset\mathfrak{B}$ 使得 $O=\bigcup\limits_{B\in\mathfrak{B}'}B$，我们证明这样的开子集族 $\mathfrak{B}\subset\tau$ 是一个拓扑基.

事实上，因 $\forall G\in\tau$, $x\in G$，由假设 $G=\bigcup\limits_{B\in\mathfrak{B}'}B$，故必有 $B\in\mathfrak{B}'\subset\mathfrak{B}$ 使得 $x\in B\subset G$，这正说明开子集族 $\mathfrak{B}\subset\tau$ 是一个拓扑基. 充分性得证.

注 从定义 3.2.5 与定理 3.2.5 可知，拓扑基 $\mathfrak{B}$是拓扑空间(X,τ)中开集族 τ 的一部分，可视为(X,τ)中的"基本"开集，(X,τ)中任何其他开集都可由一些基本开集的并来表示. 于是可将对拓扑空间(X,τ)的研究简化为对$(X,\mathfrak{B})$的研究，这样会带来许多方便.

例如，对于$\mathbb{R}=(-\infty,+\infty)$，若$(\mathbb{R},\tau)$是$\mathbb{R}$的拓扑，$\tau=\{G\subset\mathbb{R}:G$ 是$\mathbb{R}$的开子集$\}$，不难证明 $\mathfrak{B}=\{(a,b):-\infty<a<b<+\infty\}$是$\mathbb{R}$的一个拓扑基.

关于构成拓扑基的条件，有如下定理.

定理 3.2.6 设 X 是一个集合，$\mathfrak{B}$是 X 的一个子集族，若 $\mathfrak{B}$满足

(1) $\bigcup\limits_{B\in\mathfrak{B}}B=X$；

(2) $B_1,B_2\in\mathfrak{B}\Rightarrow\ \forall x\in B_1\cap B_2,\exists B\in\mathfrak{B}\ \text{s. t.}\ x\in B\subset B_1\cap B_2$，

则存在 X 上惟一的拓扑 $\tau=\left\{O\subset X:\exists\,\mathfrak{B}'\subset\mathfrak{B}\ \text{s. t.}\ O=\bigcup\limits_{B\in\mathfrak{B}'}B\right\}$，$\tau$ 以 $\mathfrak{B}$为拓扑基，使得(X,τ)成为一个拓扑空间. 反之，若 X 的一个子集族 $\mathfrak{C}$是 X 的一个拓扑基，则 $\mathfrak{C}$必满足条件(1)与(2).

证 我们证明 $\tau=\left\{O\subset X:\exists\,\mathfrak{B}'\subset\mathfrak{B}\ \text{s. t.}\ O=\bigcup\limits_{B\in\mathfrak{B}'}B\right\}$是以 $\mathfrak{B}$为拓扑基的 X 上的拓扑.

首先证明，τ 满足定义 3.2.1 的(1)、(2)、(3).

(1) $X,\varnothing\in\tau$：由定理的条件(1)，知 $X=\bigcup\limits_{B\in\mathfrak{B}}B\in\tau$，而由约定，$\varnothing\in\tau$；

(2) $A_1,A_2\in\tau\Rightarrow A_1\cap A_2\in\tau$：分两步证.

① 若 $A_1=B_1\in\mathfrak{B}$，$A_2=B_2\in\mathfrak{B}$，则 $B_1\cap B_2\in\tau$.

事实上，由定理条件(2)，得 $\forall x\in B_1\cap B_2,\exists B_x\in\mathfrak{B}$使得 $x\in B_x\subset B_1\cap B_2$. 另一方面，$B_1\cap B_2=\bigcup\limits_{x\in B_1\cap B_2}\{x\}\subset\bigcup\limits_{x\in B_1\cap B_2}B_x\subset B_1\cap B_2$，故 $B_1\cap B_2=\bigcup\limits_{x\in B_1\cap B_2}B_x$，由 τ 的组成得 $B_1\cap B_2\in\tau$.

② 当 $A_1,A_2\in\tau$ 时，则 $A_1\cap A_2\in\tau$.

因为 $A_1\in\tau\Rightarrow A_1=\bigcup\limits_{B_1\in\mathfrak{B}_1'}B_1$，$A_2\in\tau\Rightarrow A_2=\bigcup\limits_{B_2\in\mathfrak{B}_2'}B_2$. 于是，作交集 $A_1\cap A_2=\left(\bigcup\limits_{B_1\in\mathfrak{B}_1'}B_1\right)\cap\left(\bigcup\limits_{B_2\in\mathfrak{B}_2'}B_2\right)=\bigcup\limits_{\substack{B_1\in\mathfrak{B}_1'\\B_2\in\mathfrak{B}_2'}}B_1\cap B_2$. 据①，当 $B_1,B_2\in\mathfrak{B}$时，有 $B_1\cap B_2=\bigcup\limits_{x\in B_1\cap B_2}B_x\in\tau$，因此 $A_1\cap A_2=\bigcup\limits_{\substack{B_1\in\mathfrak{B}_1'\\B_2\in\mathfrak{B}_2'}}\bigcup\limits_{x\in B_1\cap B_2}B_x\in\tau$.

(3) $A_\alpha\in\tau,\alpha\in\Lambda\Rightarrow\bigcup\limits_{\alpha\in\Lambda}A_\alpha\in\tau$.

由 $A_\alpha\in\tau\Rightarrow\exists\,B_\beta^\alpha\in\mathfrak{B}_\beta'\subset\mathfrak{B}\ \text{s. t.}\ B_\alpha=\bigcup\limits_{B_\beta^\alpha\in\mathfrak{B}_\beta'}B_\beta^\alpha$. 这样，有 $\bigcup\limits_{\alpha\in\Lambda}A_\alpha=\bigcup\limits_{\alpha\in\Lambda}\bigcup\limits_{B_\beta^\alpha\in\mathfrak{B}_\beta'}B_\alpha^\beta\equiv\bigcup\limits_{B_\gamma\in\mathfrak{B}_\gamma'}B_\gamma\in\tau$. 于

是，$(X,\tau)\equiv\left(X,\left\{O\subset X: \exists\, \mathfrak{B}'\subset\mathfrak{B}\ \text{s. t.}\ O=\bigcup_{B\in\mathfrak{B}'}B\right\}\right)$成为一个拓扑空间，并且由 τ 的表示与拓扑基的定义知，$\mathfrak{B}$就是 $\tau=\left\{O\subset X: \exists\, \mathfrak{B}'\subset\mathfrak{B}\ \text{s. t.}\ O=\bigcup_{B\in\mathfrak{B}'}B\right\}$的拓扑基.

其次证明，满足定理条件(1)、(2)的集族 $\mathfrak{B}$唯一地确定一个拓扑 τ，如上面的(3)中给出的 τ 那样.

事实上，设另有一个拓扑 $\tilde{\tau}$ 以 $\mathfrak{B}$为拓扑基，根据拓扑基的定义，$\forall A\in\tilde{\tau}\Rightarrow A=\bigcup_{B\in\mathfrak{A}}B\Rightarrow A\in\tau$. 于是得到 $\tilde{\tau}\subset\tau$. 反之，由于 $\mathfrak{B}$是 $\tilde{\tau}$ 的拓扑基，故 $\mathfrak{B}\subset\tilde{\tau}$. 取 τ 中开集 $A\in\tau$，则 A 是 $\mathfrak{B}$中某些元素之并，但因 $\mathfrak{B}$是 $\tilde{\tau}$ 的拓扑基，所以 A 也是 $\tilde{\tau}$ 中某些元素之并，于是 $A\in\tilde{\tau}$，故 $\tau\subset\tilde{\tau}$. 结合二者，就证明了 $\tilde{\tau}=\tau$.

最后证明，X 的一个子集族 $\mathfrak{C}$是 X 的某个拓扑的拓扑基，则 $\mathfrak{C}$必满足定理的条件(1)与(2).

设 X 的子集族 $\mathfrak{C}$是 X 的某个拓扑 τ^* 的拓扑基，则由 $X\in\tau^*$ 知，$X=\bigcup_{C\in\mathfrak{G}\subset\mathfrak{C}}C=\bigcup_{C\in\mathfrak{C}}C$，这里 $\mathfrak{G}$ 是 $\mathfrak{C}$的某个子集族. 故 $\mathfrak{C}$满足条件(1). 又对于 $C_1,C_2\in\mathfrak{C}\subset\tau^*$，知 C_1,C_2 是 X 中的开集；因此，对于 $x\in C_1\cap C_2$，交集 $C_1\cap C_2$ 是含 x 的开集，因而是 x 的开邻域. 于是，据邻域的定义知，存在 $W_x\in\mathfrak{C}$,使得$x\in W_x\subset C_1\cap C_2$，这表明 $\mathfrak{C}$满足条件(2). 定理得证.

注 1 定理 3.2.6 的意义在于，为构造集合上的一个拓扑，只要给出满足定理条件的(1)、(2)的集族 $\mathfrak{B}$即可.

注 2 对于拓扑空间(X,τ)，设$\mathbb{S}\subset\tau$ 是 τ 的子集族，若由$\mathbb{S}$ 的有限非空子集族之交产生的子集族

$$\mathfrak{B}=\left\{B=\bigcap_{j=1}^{n}S_j: \varnothing\neq S_j\in\mathbb{S},1\leqslant j\leqslant n;\ n\in\mathbb{Z}^+\right\} \tag{3.2.2}$$

是(X,τ)的一个拓扑基，则$\mathbb{S}$ 是包含在拓扑基 $\mathfrak{B}$ 中的“子基”，称$\mathbb{S}$ 为**拓扑 τ 的子基**，或称其为**拓扑空间(X,τ)的子基**.

在一个集合上，可以赋予不同的拓扑，并且不同拓扑中的开集可以有不同的性质，下面的例子是在实数集$\mathbb{R}$ 上的三种不同的拓扑.

例 3.2.3 $\mathbb{R}$ 上的通常拓扑(也称标准拓扑)、下限拓扑与 K 拓扑.

通常拓扑(usual topology)$(\mathbb{R},\tau)$　τ 以 $\mathfrak{B}=\{(a,b): -\infty<a<b<+\infty\}$为拓扑基，用$(\mathbb{R},\mathfrak{B})$表示；显然，$\tau$ 有一个子基 $\mathbb{S}=\{\{(-\infty,a): a\in\mathbb{R}\}\cup\{(b,+\infty): b\in\mathbb{R}\}\}$.

下限拓扑(lower topology)$(\mathbb{R},\tau')$　拓扑基为 $\mathfrak{B}=\{[a,b): -\infty<a<b<+\infty\}$，用$(\mathbb{R},\mathfrak{B})$表示；

K 拓扑(K-topology)$(\mathbb{R},\tau'')$　拓扑基为 $\mathfrak{B}=\left\{(a,b)-\left\{1,\frac{1}{2},\frac{1}{3},\cdots\right\}: -\infty<a<b<+\infty\right\}$，用$(\mathbb{R},\mathfrak{B})$表示.

不难验证 $\mathfrak{B}\mathfrak{B},\mathfrak{B}$满足定理 3.2.6 中的条件，故它们都是拓扑基，从而得到$\mathbb{R}$ 上的三种不同的拓扑结构.

2）按“拓扑基的势”分类 —— 第一可数性、第二可数性

定义 3.2.6（可数性公理） 设(X,τ)是拓扑空间，若 X 有一个可数拓扑基 $\mathfrak{B}$则称(X,τ)满足**第二可数性公理（second axiom of countability）**. 若 X 中每一点 $x\in X$ 都有一个可数拓扑基 $\mathfrak{B}$，则称(X,τ)满足**第一可数性公理（first axiom of countability）**.

例 3.2.4 一维实欧氏空间$(\mathbb{R},\tau)$（即实数空间）满足第二可数性公理.

事实上，设$(\mathbb{R},\tau)$是一维拓扑空间，若 $O\subset\mathbb{R}$ 为开集，则 $\forall x\in O$，存在 $\eta>0$，使得$(x-\eta,x+\eta)\subset O$. 此时，必能取 τ 的开子集族 $\mathfrak{B}=\{B=(x-a_x,x+b_x):a_x,b_x\in\mathbb{Q}\}$中的一个子集 $B=(x-a_x,x+b_x)$，使得 $B\subset(x-\eta,x+\eta)\subset O$，这里$\mathbb{Q}$ 是有理数集. 故 $x\in B=(x-a_x,x+b_x)\subset O$，所以$\mathbb{R}$ 满足第一可数性公理.

进而，$\mathfrak{B}=\{\mathfrak{B}:x\in\mathbb{Q}\}$显然是$(\mathbb{R},\tau)$的一个可数拓扑基，故$\mathbb{R}$ 满足第二可数性公理.

不难证明，任何度量空间(X,ρ)都是满足第一可数性公理的拓扑空间. 留作习题.

显然，满足第二可数性公理的拓扑空间(X,τ)必定满足第一可数性公理，反之却未必. 而且，也有不满足第一可数性公理的拓扑空间.

2. 拓扑空间的可分性

1）拓扑空间可分性的定义

定义 3.2.7（可分性） 设(X,τ)是拓扑空间，若存在 X 的一个至多可数子集 $A=\{a_1,a_2,\cdots,a_n,\cdots\}\subset X$，使得 $\overline{A}=X$，则称 X 具有**可分性（separability）**；否则称 X 具有**不可分性（unseparability）**.

例如，欧氏空间$\mathbb{R}^n$ 是可分的，因为 $\overline{\mathbb{Q}^n}=\mathbb{R}^n$.

例 3.2.5 区间$[a,b]$上的连续函数空间 $C([a,b])$，在距离 $\rho(f,g)=\max\limits_{x\in[a,b]}|f(x)-g(x)|$之下，成为一个拓扑空间（度量空间），且它具有可分性.

事实上，多项式集 $P_0=\{p_n:p_n(x)=r_nx^n+\cdots+r_1x+r_0,x\in[a,b],r_j\in\mathbb{Q}\}$是空间 $C([a,b])$的一个可数稠密子集. 根据伯恩斯坦（Bernstein）定理，对于函数 $f\in C([a,b])$，存在有理系数多项式 $p_n(x)\in P_0$，使得 $p_n(x)$在$[a,b]$上一致收敛于 $f(x)$. 于是$\overline{P_0}=C([a,b])$. 故 $C([a,b])$具有可分性.

类似地，利用魏尔斯特拉斯（Weierstrass）定理（参看本书 7.2 节），可以证明$\overline{P_0}=L^p([a,b])(1\leqslant p<+\infty)$，故 $L^p([a,b])$也具有可分性.

当然也有不可分空间的例子.

例 3.2.6 $[0,1]$上有界函数空间$\mathbb{B}([0,1])=\{f:[0,1]\to\mathbb{R},\exists M>0,\text{s.t.}\ |f(x)|\leqslant M\}$在距离 $\rho(f,g)=\sup\limits_{x\in[0,1]}|f(x)-g(x)|$之下，成为一个拓扑空间（度量空间），它具有不可

分性.

事实上,若 $A=\{f_n\in\mathbb{B}([0,1]):n\in\mathbb{Z}^+\}$为$\mathbb{B}([0,1])$的任意可数子集,我们证明 A 在 $\mathbb{B}([0,1])$中不稠密. 只要证明存在函数 $g\in\mathbb{B}([0,1])$,但 $g\notin\overline{A}$. 由于 $[0,1]=\left[0,\frac{1}{2}\right)\cup\left[\frac{1}{2},\frac{2}{3}\right)\cup\cdots\cup\left[1-\frac{1}{n},1-\frac{1}{n+1}\right)\cup\cdots=\{1\}\cup\left(\bigcup_{n=1}^{+\infty}\left[1-\frac{1}{n},1-\frac{1}{n+1}\right)\right)$,对于 $f_n\in A,n\in\mathbb{Z}^+$,构造函数 $g:[0,1]\to\mathbb{R}$,满足

$$g(t)=\begin{cases}2, & t\in\left[1-\frac{1}{n},1-\frac{1}{n+1}\right),\forall t'\in\left[1-\frac{1}{n},1-\frac{1}{n+1}\right),|f_n(t')|\leqslant 1,\\ 0, & t\in\left[1-\frac{1}{n},1-\frac{1}{n+1}\right),\exists t''\in\left[1-\frac{1}{n},1-\frac{1}{n+1}\right),|f_n(t'')|>1,\\ 0, & t=1,\end{cases}\tag{3.2.3}$$

则 $g\in\mathbb{B}([0,1])$,且 $\forall n\in\mathbb{Z}^+,\rho(f_n,g)=\sup\limits_{x\in[0,1]}|f_n(x)-g(x)|\geqslant 1$. 因此球

$$B(g,1)=\{f\in\mathbb{B}([0,1]):\rho(f,g)<1\}$$

与集合 A 的交集是空集,于是 $g\notin\overline{A}$,故 $\overline{A}\underset{\neq}{\subset}\mathbb{B}([0,1])$. 这就证明了$\mathbb{B}([0,1])$是具有不可分性的度量空间.

2) 按拓扑空间的可分性分类 —— 可分空间、不可分空间

由拓扑空间的可分性与不可分性,我们将一个拓扑空间(X,τ)称为**可分拓扑空间(separable topological space)**,如果它具有可分性;否则,就称它为**不可分拓扑空间(unseparable topological space)**.

3.3 拓扑空间上的连续映射

3.3.1 拓扑空间之间的映射、映射的连续性

1. 连续映射

研究拓扑空间之间的连续映射,是点集拓扑学中的一个重要内容,因为连续性保持了拓扑空间的许多性质. 在两个拓扑空间(X,τ),(Y,υ)中引入连续映射,是从微积分学、实变函数理论中连续函数判定的一个充要条件得到启发的:"$f:\mathbb{R}\to\mathbb{R}$ 是连续函数,当且仅当开集 $G\subset\mathbb{R}$ 的逆像集 $f^{-1}(G)\subset\mathbb{R}$ 是开集."

定义 3.3.1(映射的连续性) 设(X,τ),(Y,υ)是两个拓扑空间. $f:X\to Y$ 是 X 到 Y 的映射. 称 f 为 X 到 Y 的**连续映射(continuous mapping)**,若对任意开集 $O\in\upsilon$,蕴含 $G=f^{-1}(O)\in\tau$;亦即,Y 中开集 $O\in\upsilon$ 的逆像集 $G=f^{-1}(O)$是 X 中的开集.

连续映射有如下等价性质.

定理 3.3.1 设(X,τ),(Y,υ)是两个拓扑空间. $f:X\to Y$ 是 X 到 Y 的映射. 则下述性质等价:

(1) f 是 X 到 Y 的连续映射;

(2) Y 中开集 $O\in\upsilon$ 的逆像集 $f^{-1}(O)$是 X 中的开集;

(3) Y 中闭集 $F\subset Y$ 的逆像集 $f^{-1}(F)\subset X$ 是 X 中的闭集;

(4) 对任一点 $x\in X$,对应点为 $y=f(x)\in Y$,则 $\forall V_y\in\upsilon, y\in V_y$, $\exists U_x\in\tau$, s. t. $f(U_x)\subset V_y$;

(5) Y 中点 $y\in Y$ 的开邻域 V 的逆像集 $f^{-1}(V)$是 X 中满足 $f(x)=y$ 的 $x\in X$ 的开邻域;

(6) $B\subset Y$ 蕴含$\overline{f^{-1}(B)}\subset f^{-1}(\overline{B})$;

(7) $A\subset X$ 蕴含 $f(\overline{A})\subset\overline{f(A)}$.

证 (1)$\Leftrightarrow$(2)由定义可得.

(2)$\Leftrightarrow$(3) 对于 Y 中任意闭集 $F\subset Y$,由定义知 $Y\backslash F\subset Y$ 为开集,即 $F^{C}=Y\backslash F\in\upsilon$. 于是,根据(2)得 $f^{-1}(Y\backslash F)\in\tau$. 故由 $f^{-1}(Y\backslash F)=f^{-1}(Y)\backslash f^{-1}(F)=X\backslash f^{-1}(F)$,且 $f^{-1}(Y\backslash F)$是 X 中的开集,故 $f^{-1}(F)$是 X 中的闭集,(3)成立; 反之亦然.

为证(2)$\Leftrightarrow$(4),我们作如下推理.

(2)$\Rightarrow$(4) 设 $x\in X, y=f(x)\in Y$. $\forall V_y\in\upsilon, y=f(x)\in V_y$,根据(2)知 $f^{-1}(V_y)$为 X 中的开集,$f^{-1}(V_y)\in\tau$,且 $x\in f^{-1}(V_y)$,从而 $f^{-1}(V_y)$是包含 x 的开集,故 $\exists U_x\in\tau$, s. t. $U_x\subset f^{-1}(V_y)$,因此 $f(U_x)\subset f(f^{-1}(V_y))\subset V_y$,此即(4).

(4)$\Rightarrow$(2) 取 $\forall O\in\upsilon$,我们证明 $f^{-1}(O)\in\tau$ 为开集. 事实上, $\forall x\in f^{-1}(O)$,对应的 $y=f(x)\in f(f^{-1}(O))\subset O$. 对此 $y\in O\in\tau$,根据(4), $\exists U_x\in\tau, x\in U_x$, s. t. $f(U_x)\subset O$,这蕴含 $x\in U_x\subset f^{-1}(f(U_x))\subset f^{-1}(O)$,故又蕴含 $f^{-1}(O)$是 X 中的开集,$f^{-1}(O)\in\tau$,此即(2).

为证(2)$\Leftrightarrow$(5),我们作如下推理.

(2)$\Rightarrow$(5) 对 $y\in Y$,有 $x\in X$,使得 $y=f(x)$. $\forall U_y\in\upsilon, y=f(x)\in U_y$,根据(2),$x\in f^{-1}(U_y)\in\tau$,从而 $f^{-1}(U_y)$为 $x\in X$ 的邻域,此即(5);

(5)$\Rightarrow$(2) $\forall O\in\upsilon$,对于每个 $y\in O$,则 O 就是 y 的一个开邻域,由(5),$f^{-1}(O)\subset X$ 是 $x\in f^{-1}(O)\subset X$ 的开邻域,这里 x 是满足 $y=f(x)$的 X 中的元. 于是,对于邻域 $f^{-1}(O)$, $\exists G\subset X, G\in\tau$, s. t. $x\in G\subset f^{-1}(O)$,故 $f^{-1}(O)$是 X 中的开集,$f^{-1}(O)\in\tau$,这就是(2).

下证(3)$\Leftrightarrow$(6).

(3)$\Rightarrow$(6) $\forall B\subset Y\Rightarrow f^{-1}(\overline{B})\subset X$ 为闭集(根据(3)) $\Rightarrow\overline{f^{-1}(B)}\subset f^{-1}(\overline{B})$(因 $B\subset\overline{B}$ 蕴含 $f^{-1}(B)\subset f^{-1}(\overline{B})$,蕴含$\overline{f^{-1}(B)}\subset\overline{f^{-1}(\overline{B})}$) $\Rightarrow$ (6)成立;

(6)$\Rightarrow$(3) 每个闭集 $F\subset Y\Rightarrow\overline{f^{-1}(F)}\subset f^{-1}(\overline{F})=f^{-1}(F)$(根据(6)) $\Rightarrow\overline{f^{-1}(F)}=f^{-1}(F)\Rightarrow f^{-1}(F)$为闭集 $\Rightarrow$ (3)成立.

最后证(6)$\Leftrightarrow$(7).

(6)$\Rightarrow$(7)　$\forall A\subset X$ 有 $f(A)\subset Y \Rightarrow f(A)\subset\overline{f(A)} \Rightarrow \overline{f^{-1}(f(A))}\subset f^{-1}(\overline{f(A)})$(由(6)) $\Rightarrow$ $f(\bar{A})\subset\overline{f(A)}$(因包含关系 $A\subset f^{-1}(f(A))$ 蕴含 $\bar{A}\subset\overline{f^{-1}(f(A))}$，蕴含 $f(\bar{A})\subset f(\overline{f^{-1}(f(A))})\subset f(f^{-1}(\overline{f(A)}))\subset\overline{f(A)}$) $\Rightarrow$ (7)成立；

(7)$\Rightarrow$(6)　$\forall B\subset Y$ 有 $f^{-1}(B)\subset X \Rightarrow$ 取(7)中的 $A\equiv f^{-1}(B)\subset X$ 得 $f(\overline{f^{-1}(B)})\subset\overline{f(f^{-1}(B))}\subset\bar{B}$(因 $f(f^{-1}(B))\subset B\subset\bar{B}$，故 $\overline{f(f^{-1}(B))}\subset\bar{B}$) $\Rightarrow f^{-1}(f(\overline{f^{-1}(B)}))\subset f^{-1}(\bar{B})$(因 $f^{-1}(\overline{f(f^{-1}(B))})=\overline{f^{-1}(B)}$) $\Rightarrow \overline{f^{-1}(B)}\subset f^{-1}(\bar{B}) \Rightarrow$ (6)成立.

在例 3.2.3 中给出了一维欧氏空间中的三种不同的拓扑，对于第一、第二种拓扑，恒同映射 $I:(\mathbb{R},\mathfrak{B})\to(\mathbb{R},\mathfrak{B})$，$I(x)=x$ 不是连续映射. 这是因为开集 $O=[a,b)\in\mathfrak{B}$的逆像集 $I^{-1}([a,b))=[a,b)$不是$(\mathbb{R},\mathfrak{B})$拓扑之下的开集.

定义 3.3.2(拓扑空间的同胚、开(闭)映射)　设(X,τ)，(Y,υ)是两个拓扑空间. 若存在 X 到 Y 的一一映射 $f:X\to Y$，使得 f 与 f^{-1} 都是连续的，亦即 f 与 f^{-1} 是**双方单值、双方连续的**，则称 f 是 X 到 Y 的**同胚(homeomorphism)映射**，并称 X 与 Y 是(拓扑)**同胚的**.

在拓扑同胚映射下保持不变的性质(或不变的量)，称为**拓扑不变性(topological invariant property)**(或称**拓扑不变量(topological invariant quantity)**).

若映射 $f:X\to Y$ 把 X 中的开集(闭集)映到 Y 中的开集(闭集)，则称 f 为 X 到 Y 的**开映射(闭映射)**.

同胚与开(闭)映射有如下的等价性定理.

定理 3.3.2　设(X,τ)，(Y,υ)是两个拓扑空间. $f:X\to Y$ 是 X 到 Y 的一一映射(单射加满射). 则下述性质等价：

(1) f 是 X 到 Y 的同胚映射；

(2) f 是 X 到 Y 的连续开映射(闭映射).

证　事实上，由 f^{-1}的连续性，$\forall U_x\in\tau$，$[(f^{-1})^{-1}](U_x)=f(U_x)\in\upsilon$，故 f 将开集映到开集，所以 f 是开映射；反之亦然. 对闭集一样推理.

定理 3.3.3　设(X,τ)，(Y,υ)，(Z,ω)是三个拓扑空间. $f:X\to Y$ 是 X 到 Y 的连续映射，$g:Y\to Z$ 是 Y 到 Z 的连续映射，则复合映射 $g\circ f:X\to Z$ 也是连续映射.

证　任取一点 $x_0\in X$，则 $y_0=f(x_0)\in Y$，$z_0=g(y_0)=g(f(x_0))\in Z$，为证复合映射 $g\circ f:X\to Z$ 的连续性，$\forall W_{z_0}\in\omega$，由 $g:Y\to Z$ 的连续性，$\exists V_{y_0}\in\upsilon$，s. t. $g(V_{y_0})\subset W_{z_0}$；再由 $f:X\to Y$ 的连续性，对于上述 $V_{y_0}\in\upsilon$，$\exists U_{x_0}\in\tau$，s. t. $f(U_{x_0})\subset V_{y_0}$.

于是，$\forall W_{z_0}\in\omega$，$\exists U_{x_0}\in\tau$，s. t. $g(f(U_{x_0}))\subset g(V_{y_0})\subset W_{z_0}$，这就是 $g\circ f:X\to Z$ 在点 x_0 的连续性. 再由 $x_0\in X$ 的任意性，得到复合映射 $g\circ f:X\to Z$ 的连续性.

例 3.3.1　设(X,τ_1)，(X,τ_2)为同一集合 X 上的两个拓扑结构. $I:(X,\tau_1)\to(X,\tau_2)$为恒同映射，$I(x)=x$，$x\in X$. 则 $I:X\to X$ 为连续映射，当且仅当 $\tau_2\subset\tau_1$.

证 由连续映射的充要条件(定理 3.3.1(2)),$\forall O\in\tau_2$,$\exists G\in\tau_1$, s.t. $I(G)\subset O$,但 $I(G)=G$,故 $G\subset O$ 表明 $O\in\tau_1$,此即 $\tau_2\subset\tau_1$.反之亦然.

2. 极限

在微积分学中首先定义“极限”,而后才定义“连续性”.但是,在拓扑空间中,却是先引入连续概念.然后引入极限概念.

在拓扑空间中引进极限有很多种方法,用“格网”、“滤子”等,但是都较为复杂.我们采用较直接的方法 —— 序列的极限.

定义 3.3.3(序列、极限) 设(X,τ)为拓扑空间.每个映射 $S:j\in\mathbb{Z}^+\to x_j\in X$ 称为 X 中的一个**序列**(**sequence**),也称为**点列**,记为$\{x_j\}_{j\in\mathbb{Z}^+}$,简记为$\{x_j\}$.

若存在 $x\in X$,使得对包含 x 的每个开集 $G\in\tau$,$\exists N\in\mathbb{Z}^+$,当 $j>N$ 时,有 $x_j\in G$,则称 x 是序列$\{x_j\}_{j\in\mathbb{Z}^+}$的**极限**(**limit**),也称**序列**$\{x_j\}_{j\in\mathbb{Z}^+}$**收敛于点** x,记为 $\lim\limits_{j\to+\infty}x_j=x$.

定义 3.3.4(子列) 设(X,τ)为拓扑空间.$S:\mathbb{Z}^+\to X$ 与 $S_1:\mathbb{Z}^+\to X$ 是(X,τ)中的两个序列$\{x_j\}_{j\in\mathbb{Z}^+}$与$\{x_k\}_{k\in\mathbb{Z}^+}$,若存在严格递增映射 $W:j\in\mathbb{Z}^+\to k=W(j)\in\mathbb{Z}^+$(即 $\forall j_1,j_2\in\mathbb{Z}^+$,若 $j_1<j_2$,则 $k_1=W(j_1)<k_2=W(j_2)$),使得 $x_k=x_{W(j)}$,则称**序列**$\{x_k\}_{k\in\mathbb{Z}^+}$**是序列**$\{x_j\}_{j\in\mathbb{Z}^+}$**的子列**(**subsequence**).亦即,$\forall j\in\mathbb{Z}^+\Rightarrow W(j)\in\mathbb{Z}^+$,$x_{W(j)}\in\{x_j\}_{j\in\mathbb{Z}^+}$.换言之,子列$\{x_{W(j)}\}_{j\in\mathbb{Z}^+}$中的第 $j\in\mathbb{Z}^+$ 个点,恰好是原来序列$\{x_j\}_{j\in\mathbb{Z}^+}$中的第 $W(j)\in\mathbb{Z}^+$ 个点.子列也常记为$\{x_{k_j}\}_{j\in\mathbb{Z}^+}$,简记为$\{x_{k_j}\}$.

定理 3.3.4 设(X,τ)是拓扑空间,$\{x_j\}$是 X 中的一个序列.则

(1) 若$\{x_j\}$是常值序列,亦即 $\forall x_j=x\in X$,则 $\lim\limits_{j\to+\infty}x_j=\lim\limits_{j\to+\infty}x=x$;

(2) 若$\{x_j\}$收敛于 $x\in X$,则其任一个子列$\{x_{k_j}\}$也收敛于 $x\in X$;

(3) 设(X,τ),(Y,υ)是两个拓扑空间,$f:X\to Y$ 是 X 到 Y 的连续映射.若 X 中的序列$\{x_j\}$收敛到 $x\in X$,则 Y 中的序列$\{f(x_j)\}$必收敛到 $f(x)\in Y$.

证 (1)、(2)是显然的.

为证(3),即证 $\lim\limits_{j\to+\infty}x_j=x$ 蕴含 $\lim\limits_{j\to+\infty}f(x_j)=f(x)$.

对于 $y=f(x)$,$x\in X$,取 $\forall V_y\in\upsilon$ 为 $y=f(x)$的开邻域,因 $f:X\to Y$ 为连续映射,故 $\exists U_x\in\tau$,使得 $f(U_x)\subset V_y$.另一方面,由 $\lim\limits_{j\to+\infty}x_j=x\in U_x$,$\exists N\in\mathbb{Z}^+$,s.t. 当 $j>N$ 时,有 $x_n\in U_x$.于是,当 $j>N$ 时,$f(x_j)\in f(U_x)\subset V_y$,这就是 $\lim\limits_{j\to+\infty}f(x_j)=f(x)$.

注 1 在微积分学中,上述定理中的(2)与(3)的关系是充分必要的.例如,我们有(3)的逆命题:设(X,ρ)为度量空间,(Y,υ)为拓扑空间,$f:X\to Y$ 是 X 到 Y 的映射.则 f 为连续映射的充分必要条件是 X 中的序列$\{x_j\}$收敛到 $x\in X$,蕴含 Y 中的序列$\{f(x_j)\}$收敛到 $f(x)\in Y$.但在一般拓扑空间中,其逆却未必成立.这就是在一般拓扑空间中,不用序列的极限来定

义连续性的原因.

注 2 对于拓扑空间(X,τ)中的一个序列$\{x_j\}$,它可以有极限,也可以没有;如果有极限,也未必是惟一的.

例如,设(X,τ)为拓扑空间,其中至少含两个点.取$\tau=\{X,\varnothing\}$,亦即其开集只有X与$\varnothing$.于是,在此拓扑下,$\forall\{x_n\}\subset X$,对任一个元$x\in X$,都有$\lim\limits_{j\to+\infty}x_j=x$.由此可见,在拓扑$\tau=\{X,\varnothing\}$之下,任一个序列可以有两个不同的极限.

3.3.2 拓扑空间的子空间、积空间、商空间

1. 拓扑的精粗

定义 3.3.5(拓扑的精粗) 在一集合X上给出两个拓扑τ_1与τ_2,构成两个拓扑空间(X,τ_1)与(X,τ_2).若$\tau_1\subset\tau_2$,则称**由τ_1决定的拓扑比由τ_2决定的拓扑粗(coarse)**(或**弱(weak)**),或**由τ_2决定的拓扑比由τ_1决定的拓扑精(fine)**(或**强(strong)**).

例如,(X,τ_1),$\tau_1=\{X,\varnothing\}$是X上的最粗拓扑(或称最弱拓扑);

(X,τ_2),$\tau_2=\mathfrak{P}(X)=\{A:A\subset X\}$是$X$上的最精拓扑(或称最强拓扑).这两种拓扑统称为$X$上的**平凡拓扑(trivial topology)**.

以$\mathfrak{B}=\{(a,b):-\infty<a<b<+\infty\}$为拓扑基的$\mathbb{R}$上的通常拓扑$(\mathbb{R},\tau)$,就是比$\mathbb{R}$上的最粗拓扑要精、比最精拓扑要粗的**非平凡拓扑(non-trivial topology)**.

2. 子空间

定义 3.3.6(拓扑空间的子空间) 设(X,τ)为拓扑空间,X_0为X的子集,取映射$I:X_0\to X$为恒同映射,即$x=I(x)$,$x\in X_0\subset X$,令

$$\tau_0=\{G\subset X_0:G=O\cap X_0,\forall O\in\tau\},\tag{3.3.1}$$

把τ_0定义为X_0中的开集族,则(X_0,τ_0)决定了X_0的一个拓扑(不难验证τ_0满足定义3.2.1中的(1)~(3)).称拓扑τ_0为(X,τ)在子集X_0上定义的**相对拓扑(relative topology)**,或**子拓扑**,记为(X_0,τ_0),并称(X_0,τ_0)为(X,τ)的**子拓扑空间**,简称**子空间**,如图3.3.1所示.称恒同映射$I:X_0\to X$为嵌入映射.

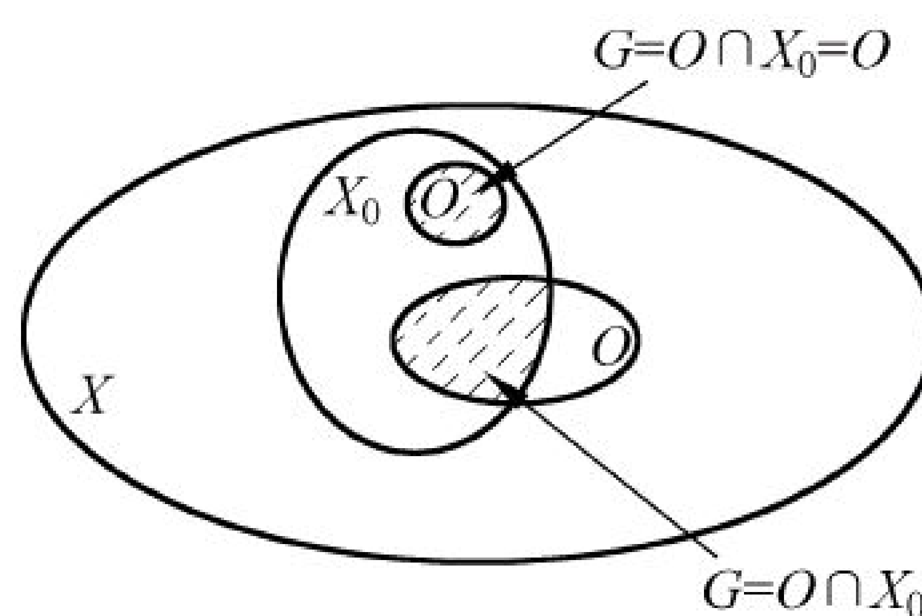


图 3.3.1 子空间

显然,若$\mathfrak{B}$是(X,τ)拓扑基,则$\mathfrak{B}=\{B\cap X_0:B\in\mathfrak{B}$是$(X_0,\tau_0)$的拓扑基.

例 3.3.2 $\mathbb{R}$,$\mathbb{R}^2$是$\mathbb{R}^3$的子空间,对于$(\mathbb{R}^3,\tau_3)$,取拓扑基

$$\mathfrak{B}=\{G:G=\{(x_1,x_2,x_3)\in\mathbb{R}^3:\sqrt{x_1^2+x_2^2+x_3^2}<q\},q\in\mathbb{Q}\};$$

对于$(\mathbb{R}^2,\tau_2)$,有$\mathbb{R}^2$中的子拓扑基

$$\mathfrak{B}=\{G:G=\{(x_1,x_2,0)\in\mathbb{R}^2:\sqrt{x_1^2+x_2^2}<q\},q\in\mathbb{Q}\};$$

对于$(\mathbb{R}^1,\tau_1)$，则有$\mathbb{R}$中的子拓扑基

$$\mathfrak{B}=\{G:G=\{(x_1,0,0)\in\mathbb{R}^1:|x_1|<q\},q\in\mathbb{Q}\}.$$

下面给出子拓扑的结构定理.

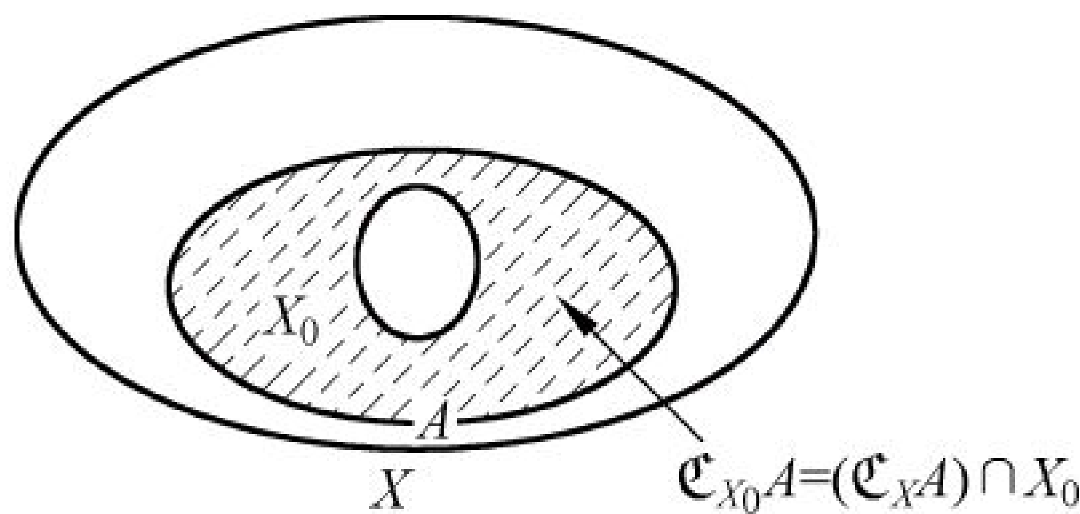


图 3.3.2 补集

定理 3.3.5 设(X,τ)为拓扑空间，$X_0\subset X$为X的子集，则对于子空间(X_0,τ_0)，有

(1) $G\subset X_0$是X_0中的开集$\Leftrightarrow$存在X的开集$O\in\tau$，使得$G=O\cap X_0$；

(2) $C\subset X_0$是X_0中的闭集$\Leftrightarrow$存在X的闭集$F\subset X$，使得$C=F\cap X_0$；

(3) 若$\bar{C}_{X_0}$是X_0的子集$C\subset X_0$关于X_0的闭包，则$\bar{C}_{X_0}=\bar{C}_X\cap X_0$，这里$\bar{C}_X$是$C$关于$X$的闭包；

(4) 子集$A\subset X_0$满足$\complement_{X_0}A=(\complement_X A)\cap X_0$，亦即，子集$A\subset X_0$关于$X_0$补集$X_0\backslash A=\complement_{X_0}A$等于它关于$X$的补集$X\backslash A=\complement_X A$与$X_0$的交.

证明留作习题.

例 3.3.3 设$(\mathbb{R},\tau)$为一维欧氏空间，试确定$\mathbb{R}$的子集$X_0=[-1,1]$的子拓扑.

取$(\mathbb{R},\tau)$的拓扑基$\mathfrak{B}=\{(a,b):-\infty<a<b<+\infty\}$，于是$X_0=[-1,1]$的拓扑基可取为$\mathfrak{B}_{[-1,1]}=\{G:G=(a,b)\cap[-1,1],-\infty<a<b<+\infty\}$. $\mathfrak{B}_{[-1,1]}$中的集合呈$(a,b)\cap[-1,1]$形式，具体为

$$(a,b)\cap[-1,1]=\begin{cases}(a,b), & a\in(-1,1),b\in(-1,1),\\ [-1,b), & a\notin[-1,1],b\in(-1,1),\\ (a,1], & a\in(-1,1),b\notin[-1,1],\\ [-1,1], & a,b\notin[-1,1],(a,b)\cap[-1,1]\neq\varnothing,\\ \varnothing, & a,b\notin[-1,1],(a,b)\cap[-1,1]=\varnothing.\end{cases}\tag{3.3.2}$$

故子空间$(X_0,\tau_0)=([-1,1],\tau_0)$中的开集与空间$\mathbb{R}$中的开集是有区别的. 但是，若集合$G\subset X_0$，则$G\in\tau\Leftrightarrow G\in\tau_0$.

下面是嵌入映射的连续性定理.

定理 3.3.6 设(X,τ)是拓扑空间，(X_0,τ_0)是(X,τ)的子空间. 则嵌入映射

$$I:X_0\to X,\quad I(x)=x,\quad x\in X_0\subset X\tag{3.3.3}$$

是X_0到X的一对一的连续映射；进而，由子拓扑决定的$X_0\subset X$的拓扑，是X_0中使得嵌入映射为连续映射的最粗拓扑.

证 (X_0,τ_0)的子拓扑是$\tau_0=\{G\subset X_0:G=O\cap X_0,\forall O\in\tau\}$，于是，由于嵌入映射$I:X_0\to X$满足$x\in X_0\Rightarrow I(x)\in X$，故$\forall O\in\tau$，根据嵌入映射的定义，$I^{-1}(O)=O\cap X_0$，由此得到$I^{-1}(O)$是$X_0$中的开集，$I^{-1}(O)\in\tau_0$. 这蕴含嵌入映射$I:X_0\to X$是$X_0$到$X$的连续

映射.

进而,若(X_0,υ)是子集X_0上的另一个拓扑,使得嵌入映射$I:(X_0,\upsilon)\to(X,\tau)$连续,则$\upsilon$至少要包含$\tau_0=\{G\subset X_0:G=O\cap X_0,O\in\tau\}$中的开集,才能保持嵌入映射$I$的连续性,故$\tau_0\subset\upsilon$,这表明此$\tau_0$是使得$I:X_0\to X$连续的最粗拓扑.

3. 积空间

定义 3.3.7(拓扑空间的积空间) 设(X,τ),(Y,υ)为两个拓扑空间,定义X与Y的积集为$X\times Y=\{(x,y):x\in X,y\in Y\}$. 赋予其拓扑基

$$\mathfrak{B}_{\times\upsilon}=\{G\times O:G\in\tau,O\in\upsilon\},\tag{3.3.4}$$

亦即$X\times Y$的拓扑基$\mathfrak{B}_{\times\upsilon}$中的子集定义为$\tau$中开子集$G\in\tau$与$\upsilon$中开子集$O\in\upsilon$的积集$G\times O$.

我们验证,$\mathfrak{B}_{\times\upsilon}$满足定理3.2.6中的(1)、(2). 事实上,

(1) 取$G=X\in\tau$, $O=Y\in\upsilon$,故$X\times Y=G\times O\in\mathfrak{B}_{\times\upsilon}$;

(2) 若$A=G_1\times O_1\in\mathfrak{B}_{\times\upsilon}$,$B=G_2\times O_2\in\mathfrak{B}_{\times\upsilon}$,则

$$A=G_1\times O_1,\quad G_1\in\tau,O_1\in\upsilon;\quad B=G_2\times O_2,\quad G_2\in\tau,O_2\in\upsilon.$$

于是

$$A\cap B=(G_1\times O_1)\cap(G_2\times O_2)=(G_1\cap G_2)\times(O_1\cap O_2)\in\mathfrak{B}_{\times\upsilon}.$$

这样,在积集$X\times Y$上赋予拓扑$\tau\times\upsilon$,它以(3.3.4)式中的$\mathfrak{B}_{\times\upsilon}$为拓扑基,称$\tau\times\upsilon$为$\tau$与$\upsilon$的**积拓扑(product topology)**,而称$X\times Y$**为X与Y的积拓扑空间(product topological space)**,简称**积空间**,记为$(X\times Y,\tau\times\upsilon)$.

例 3.3.4 $\mathbb{R}^3$是$\mathbb{R}^2$与$\mathbb{R}$的直积集,$\mathbb{R}^3=\mathbb{R}^2\times\mathbb{R}$,$\mathbb{R}^3$的拓扑可视为积空间$\mathbb{R}^2\times\mathbb{R}$上的积拓扑.

注 上面积空间的定义可以推广到有限多个拓扑空间的积空间情形,即

$$(X_1,\tau_1),\cdots,(X_k,\tau_k)\Rightarrow(X,\tau)=(X_1\times\cdots\times X_k,\tau_1\times\cdots\times\tau_k).$$

定义 3.3.8(投影映射) 设$(X_1,\tau_1),\cdots,(X_k,\tau_k)$为$k$个拓扑空间,积空间为$(X,\tau)=(X_1\times\cdots\times X_k,\tau_1\times\cdots\times\tau_k)$. 称映射

$$\mathrm{Pr}_j:X\to X_j,\quad \mathrm{Pr}_j(x_1,x_2,\cdots,x_k)=x_j,\quad j=1,2,\cdots,k\tag{3.3.5}$$

为X到$X_j(j=1,2,\cdots,k)$的**投影映射**,其中$(x_1,x_2,\cdots,x_k)\in X_1\times X_2\times\cdots\times X_k$. 称$X_j(j=1,2,\cdots,k)$为$X$的**投影空间(projective space)**.

下面是投影映射的连续性定理.

定理 3.3.7 设(X,τ)是积拓扑空间$X=X_1\times X_2\times\cdots\times X_k$,则(3.3.5)式中的投影映射$\mathrm{Pr}_j:X\to X_j(j=1,2,\cdots,k)$是$X$到$X_j(j=1,2,\cdots,k)$的连续的、满的、开映射;进而,这个积拓扑是使得每个投影映射Pr_j为连续的最粗拓扑.

证 只需对 $k=2$ 证明. 设 (X,τ) 为积拓扑空间 $(X,\tau)=(X_1\times X_2,\tau_1\times\tau_2)$, $\mathrm{Pr}_1:(x_1,x_2)\to x_1$, $\mathrm{Pr}_2:(x_1,x_2)\to x_2$. 只对 $\mathrm{Pr}_1:(x_1,x_2)\to x_1$ 证明即可.

$\mathrm{Pr}_1:(x_1,x_2)\to x_1$ 连续 任取 $G_1\in\tau_1$, 则 $(\mathrm{Pr}_1)^{-1}(G_1)=G_1\times X_2$ 是 (X,τ) 中的开集, 故对于投影映射 Pr_1 而言, 开集 $G_1\in\tau_1$ 的逆像集 $(\mathrm{Pr}_1)^{-1}(G_1)\in\tau$ 是开集, 因此 Pr_1 连续.

$\mathrm{Pr}_1:(x_1,x_2)\to x_1$ 是满射 对于任意 $x_1\in X_1$, 必有 $(x_1,x_2)\in X=X_1\times X_2$, 使得 $\mathrm{Pr}_1(x_1,x_2)=x_1$, 所以 Pr_1 是满射.

$\mathrm{Pr}_1:(x_1,x_2)\to x_1$ 是开映射 任取 $O=G_1\times G_2\in\tau=\tau_1\times\tau_2$, 则 $\mathrm{Pr}_1(O)=G_1$, 故它是开映射.

$(X,\tau)=(X_1\times X_2,\tau_1\times\tau_2)$ 是使得 Pr_1 连续的最粗拓扑 事实上, 由 Pr_1 在拓扑 $(X_1\times X_2,\tau_1\times\tau_2)$ 之下是连续的, 故 $\forall G_1\in\tau_1 \Rightarrow (\mathrm{Pr}_1)^{-1}(G_1)\in\tau$. 设有另一拓扑 $(X,\tilde{\tau})$, 使得 Pr_1 连续, 则对于上述 $G_1\in\tau_1$, 也必有 $(\mathrm{Pr}_1)^{-1}(G_1)\in\tilde{\tau}$. 因此 $\tau\subset\tilde{\tau}$. 故 $(X,\tau)=(X_1\times X_2,\tau_1\times\tau_2)$ 是使得投影映射 Pr_1 连续的最粗拓扑. 定理得证.

4. 商空间

定义 3.3.9(拓扑空间的商空间) 设 (X,τ) 是拓扑空间. 在 X 中的任意两元素之间定义一个等价关系～: 称 $x,y\in X$ 按等价关系彼此等价, 记为 $x\sim y$, 满足

(1) 自反性 $x\sim x$;

(2) 对称性 $x\sim y \Leftrightarrow y\sim x$;

(3) 传递性 $x\sim y, y\sim z \Rightarrow x\sim z$.

将 X 中的元按等价关系～分为等价类, 记 $X_a=\{x\in X: x\sim a\}$, 称 X_a 为 $a\in X$ 的**等价类**, 也记为 $X_a\equiv[a]$. 记等价类的全体为

$$\widetilde{X}=\{X_a: a\in X\}=\{[a]: a\in X\}, \tag{3.3.6}$$

并称由等价类所成的**集 $\widetilde{X}$ 为 X 关于等价关系～的商集**, 记为 $\widetilde{X}=X/\sim$. 令映射

$$\pi: a\to[a] \tag{3.3.7}$$

为由空间 X 到商集 $\widetilde{X}$ 上的映射, 满足 $\pi(a)=[a]$, 称为**商映射**.

赋予商集 $\widetilde{X}$ 以拓扑

$$\tilde{\tau}=\{O\subset\widetilde{X}: \pi^{-1}(O)\in\tau\}, \tag{3.3.8}$$

亦即, $\widetilde{X}$ 中的开集 $O\subset\widetilde{X}$ 的原像集 $\pi^{-1}(O)$ 是 τ 中的开集. 这样, 使得 $\widetilde{X}$ 成为拓扑空间, 称为 **X 关于等价关系～的商拓扑空间**, 简称**商空间**.

下面验证, (3.3.8) 式中的拓扑 $\tilde{\tau}$ 满足定义 3.2.1 中的三个条件.

事实上, 由商映射 $\pi: a\to[a]$ 的定义, 得

(1) $\pi^{-1}(\widetilde{X})=X\in\tau,\pi^{-1}(\varnothing)=\varnothing\in\tau$,故$\widetilde{X},\varnothing\in\tilde{\tau}$;

(2) $\forall O_1,O_2\in\tilde{\tau}\Rightarrow\pi^{-1}(O_1),\pi^{-1}(O_2)\in\tau\Rightarrow\pi^{-1}(O_1)\cap\pi^{-1}(O_2)\in\tau$
$\Rightarrow\pi^{-1}(O_1\cap O_2)=\pi^{-1}(O_1)\cap\pi^{-1}(O_2)\in\tau\Rightarrow O_1\cap O_2\in\tilde{\tau}$;

(3) $\forall O_\alpha\in\tilde{\tau},\alpha\in\Lambda\Rightarrow\pi^{-1}(O_\alpha)\in\tau,\alpha\in\Lambda\Rightarrow\bigcup_{\alpha\in\Lambda}\pi^{-1}(O_\alpha)\in\tau$

$$\Rightarrow\pi^{-1}\left(\bigcup_{\alpha\in\Lambda}O_\alpha\right)=\bigcup_{\alpha\in\Lambda}\pi^{-1}(O_\alpha)\in\tau\Rightarrow\bigcup_{\alpha\in\Lambda}O_\alpha\in\tilde{\tau}.$$

验证完毕.

对于实数集$\mathbb{R}$,若我们只需研究有理数与无理数,就可以给出等价关系

$$\sim=\{(x,y)\in\mathbb{R}^2: x,y\in\mathbb{Q},\text{或 }x,y\notin\mathbb{Q}\},$$

亦即,$x,y\in\mathbb{R},x\sim y\Leftrightarrow x,y\in\mathbb{Q}$,或$x,y\notin\mathbb{Q}$,在这一等价关系下,所有有理数都彼此等价,所有无理数也彼此等价.这样,实数集$\mathbb{R}$就被分为两个等价类,商集$\mathbb{R}/\sim$中只有两个等价类,一类是有理数集类,记为$[r]$;另一类是无理数集类,记为$[S]$.赋予商集$\mathbb{R}/\sim$以商拓扑$\tau_{\mathbb{R}/\sim}=\{\mathbb{R}/\sim,\varnothing,[r],[s]\}$,就得到商空间$(\mathbb{R}/\sim,\tau_{\mathbb{R}/\sim})$.

下面给出商映射的连续性定理.

定理 3.3.8 设$(\widetilde{X},\tilde{\tau})$是拓扑空间$(X,\tau)$在等价关系$\sim$之下的商拓扑空间,则商映射$\pi:X\to\widetilde{X}$是$X$到$\widetilde{X}$的连续的满映射;进而,商拓扑是使得商映射为连续映射的最精拓扑.

证 (1) 商映射$\pi:X\to\widetilde{X}$是满的连续映射 $\forall[x]\in\widetilde{X}=X/\sim$,必(至少)有一个$x\in X$,使得$\pi(x)=[x]$,故$\pi:X\to\widetilde{X}$为满射;进而,由商拓扑的定义,$\forall U\in\tilde{\tau}$,有$\pi^{-1}(U)\in\tau$,说明$\tilde{\tau}$中开集的逆像集是$\tau$中的开集,因此映射$\pi:X\to\widetilde{X}$是连续的.

(2) 商拓扑$\tilde{\tau}$是使得商映射$\pi:X\to\widetilde{X}$连续的最精拓扑 设$\widetilde{X}$上另有一个拓扑$(\widetilde{X},\upsilon)$,使得商映射$\pi:X\to\widetilde{X}$连续,于是,$\forall A\in\upsilon\Rightarrow\pi^{-1}(A)\in\tau$.然而,商拓扑$\tilde{\tau}$已经含有所有使得$\pi^{-1}(O)\in\tau$的"开集"$O\subset\widetilde{X}$,因此满足$A\in\upsilon\Rightarrow\pi^{-1}(A)\in\tau$的$A\subset\widetilde{X}$也必有$A\in\tilde{\tau}$,故$\upsilon\subset\tilde{\tau}$.于是$\widetilde{X}$上任何使得商映射$\pi$连续的其他拓扑$\upsilon$都必定包含在$\tilde{\tau}$中,$\upsilon\subset\tilde{\tau}$.因此,商拓扑$\tilde{\tau}$是使得商映射$\pi$连续的最精拓扑.定理得证.

例 3.3.5 设$p\in\mathbb{Z}^+$为一个素数,令$p\mathbb{Z}=\{\cdots,-2p,-p,0,p,2p,\cdots\}$.在$\mathbb{Z}$中定义等价关系$\sim$: $a,b\in\mathbb{Z},a\sim b\Leftrightarrow a-b\in p\mathbb{Z}$.于是

$$\mathbb{Z}/\sim=\{0,1,\cdots,p-1\}.\tag{3.3.9}$$

例 3.3.6 设单位正方形为$I^2=[0,1]\times[0,1]$,定义等价关系

$$\sim=\{(x,y)\in I^2\times I^2: x=y,\text{或 }x_1,y_1\in\{0,1\},x_2=y_2\},\tag{3.3.10}$$

下面确定商空间$I^2/\sim$.

事实上,$I^2=[0,1]\times[0,1]$中的上述等价关系是:对于$x=(x_1,x_2),y=(y_1,y_2)\in I^2$,

$$x = y \quad 或 \quad x_1, y_1 \in \{0,1\}, \quad x_2 = y_2.$$

于是，分别在对边 $OB(x_1,y_1=0)$，$AC(x_1,y_1=1)$上的对应点 $\alpha,\beta(x_2=y_2)$同属一个等价类，以及不在 OB,AC 上的 $x\in I^2$，其等价类就是$[x]=\{x\}$. 故商集 $I^2/\sim$ 中的元可视为将 I^2 的一对竖直的对边 OB,AC 上每一对纵坐标相同的点$(0,y)$与$(1,y)$相粘合所得到的圆圈. 赋予其商拓扑就成为商空间 $I^2/\sim$. 可以证明，商空间 $I^2/\sim$ 与一截管子(圆柱面的一段)

$$\{(x_1,x_2,x_3)\in\mathbb{R}^3 : x_1^2+x_2^2=1, 0\leqslant x_3\leqslant 1\} = S^1\times I$$

同胚(图 3.3.3). 考虑映射 $f:I^2\to S^1\times I\subset\mathbb{R}^3$，由$(t,s)\to f(t,s)=((\cos2\pi t,\sin2\pi t),s)$，其中$(t,s)\in I^2$，$f(t,s)=((\cos2\pi t,\sin2\pi t),s)\in S^1$. 此 $f:I^2\to S^1\times I\subset\mathbb{R}^3$ 显然是连续满射. 不难证明，存在 $I^2/\sim$到 $S^1\times I$ 的同胚映射 $f^*:I^2/\sim\to S^1\times I$，使得 $f=f^*\circ\pi:I^2\to S^1\times I$.

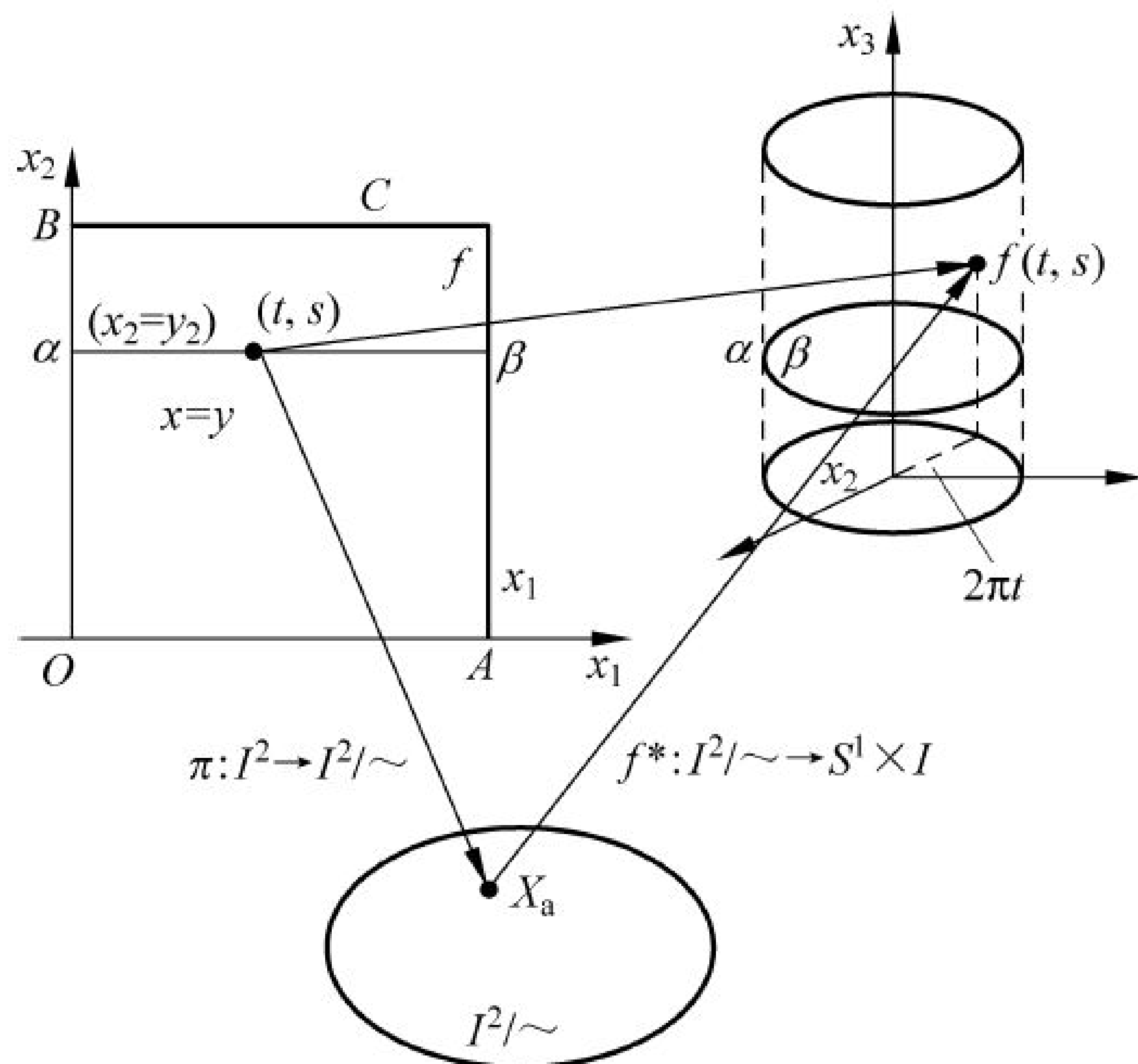


图 3.3.3

用上面的等价关系，将 I^2 的一对竖直的对边 OB,AC 上每一对对应点$(0,y)$、$(1,1-y)$相粘合，就得到 Möbius 带.

3.4 拓扑空间的重要性质

3.4.1 拓扑空间的分离性

分离性在研究拓扑空间中的极限、连续性、紧性等问题中起重要作用，而且也是拓扑空间分类的一种原则. 本小节给出拓扑空间的各种可分性与各种分离性公理.

定义 3.4.1(点的可分性) 设(X,τ)是拓扑空间. 对于任意两点 $x,y\in X$，$x\neq y$，称

(1) x 与 y 是**弱可分的(weakly separable)**，若至少存在开集 $G\in\tau$，使得 $x\in G$，但 $y\notin G$，

见图 3.4.1；

图 3.4.1　弱可分性

(2) x 与 y 是**可分的(separable)**，若存在两个开集 $G_1,G_2\in\tau$，使得 $x\in G_1,y\in G_2$，但 $x\notin G_2,y\notin G_1$，见图 3.4.2；

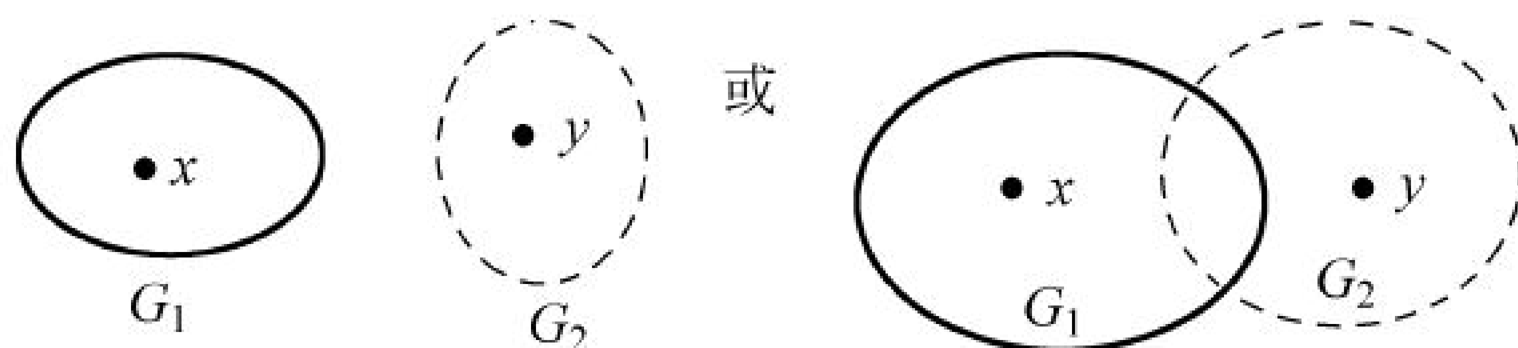


图 3.4.2　可分性

(3) x 与 y 是**分离的(separating)**，若存在两个开集 $G_1,G_2\in\tau$，使得 $x\in G_1,y\in G_2$，并且 $G_1\cap G_2=\varnothing$，见图 3.4.3.

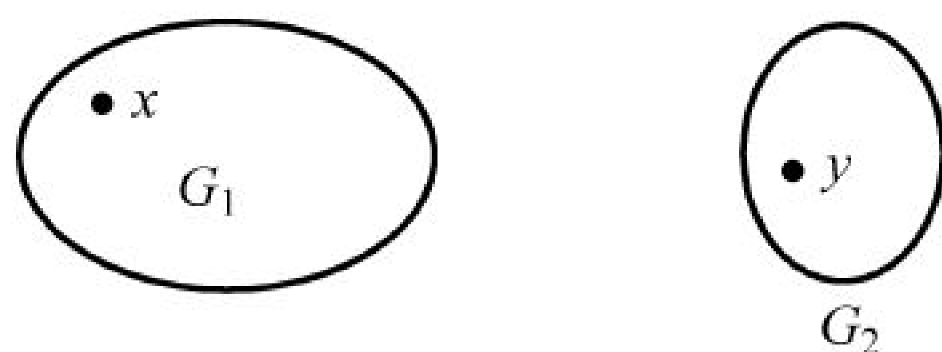


图 3.4.3　分离性

类似地，可定义**两个互不相交的集合** $A\subset X,B\subset X$ **的弱可分、可分、分离性.**

定义 3.4.2(拓扑空间的分离性公理)　设 (X,τ) 是拓扑空间.

(1) T_0 公理(T_0 axiom)　若 X 中任意两点都是弱可分的，则称 X 满足 T_0 公理，并称 X 是 T_0 型拓扑空间，简称 T_0 空间；

(2) T_1 公理(T_1 axiom)　若 X 中任意两点都是可分的，则称 X 满足 T_1 公理，并称 X 是 T_1 型拓扑空间，简称 T_1 空间；

(3) T_2 公理(T_2 axiom)　若 X 中任意两点 $x,y\in X,x\neq y$ 都是分离的，则称 X 满足 T_2 公理，并称 X 是 T_2 型拓扑空间，简称 T_2 空间，通常称为 **Hausdorff 空间**；

(4) 正则公理(regular axiom)　若 X 中任一点与不含此点的非空闭集都是分离的，则称 X 满足正则公理，并称 X 是正则拓扑空间，简称**正则空间(regular space)**；

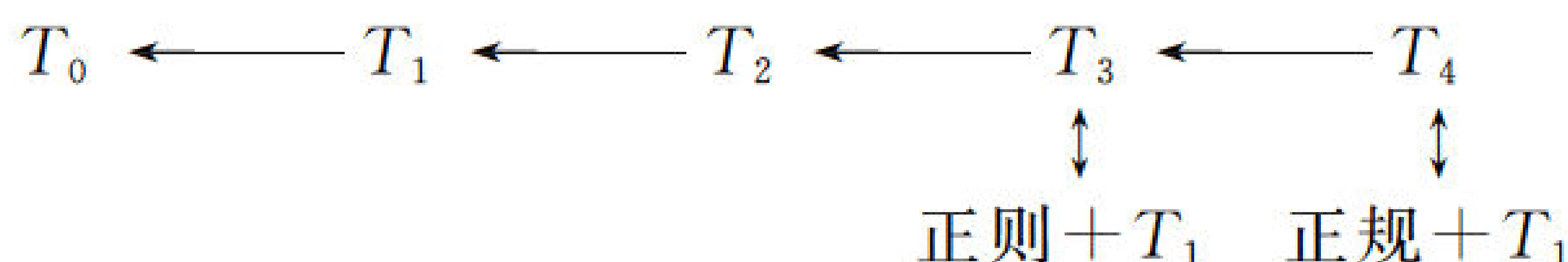
(5) 正规公理(normal axiom)　若 X 中任意两个互不相交的非空闭集都是分离的，则称 X 满足正规公理，并称 X 是**正规空间(normal space)**；

(6) T_3 公理(T_3 axiom)　若 X 是满足正则公理的 T_1 型拓扑空间，则称 X 满足 T_3 公

理，并称 X 是 T_3 型拓扑空间，简称 T_3 空间；

(7) T_4 公理(T_4 axiom) 若 X 是满足正规公理的 T_1 拓扑空间，则称 X 满足 T_4 公理，并称 X 是 T_4 型拓扑空间，简称 T_4 空间.

各分离性公理间的关系可用下图表示：



例 3.4.1(各种空间的例子) 设(X,τ)为一个拓扑空间.

非 T_0 拓扑 若 X 中至少含两个点，$\tau=\{X,\varnothing\}$为最粗拓扑，则(X,τ)不是 T_0 空间.

非 T_1 的 T_0 拓扑 若 $X=\{a,b:a\neq b\}$，$\tau=\{X,\varnothing,\{a\}\}$，则$(X,\tau)$是 T_0 空间，但非 T_1.

非 T_2 的 T_1 拓扑 若 $X=\{a_1,a_2,\cdots,a_n,\cdots\}$为可数集，取**余有限拓扑**

$$\tau=\{X,\varnothing,\{O\subset X:\complement O \text{ 为有限子集}\}\},$$

则(X,τ)是 T_1 空间，但非 T_2.

非 T_0 的正则且正规拓扑 若 $X=\{1,2,3\}$，$\tau=\{X,\varnothing,\{1\},\{2,3\}\}$，则$(X,\tau)$正则且正规，但非 T_0，也非 T_1,T_2.

非正则的正规拓扑 若 $X=\{1,2,3\}$，$\tau=\{X,\varnothing,\{1\},\{2\},\{1,2\}\}$，则$(X,\tau)$正规，但非正则.

非正规的正则拓扑 若 $X=\{x=(x_1,x_2)\in\mathbb{R}^2:x_2\geqslant 0\}$，

$$\tau=\{X,\varnothing,\{B(x,\varepsilon):x\in X,\varepsilon\in(0,x_2)\}\cup\{B(x,x_2)\cup\{(x_1,0)\}:x\in X\}\},$$

则(X,τ)正则，但非正规.

非正规、非正则的 T_2 拓扑 若 $X=\mathbb{R}$，$K=\left\{\dfrac{1}{n}:n\in\mathbb{Z}^+\right\}$，$\tau=\{G\backslash A:A\in\tau_u\}$，其中 τ_u 为 $\mathbb{R}$ 上的通常拓扑，则(X,τ)为非正则、非正规的 T_2 拓扑.

度量空间、赋范线性空间、内积空间、欧氏空间等，都是 T_4 型拓扑空间.

下面列举关于拓扑空间型的表征的几个重要定理.

定理 3.4.1(T_0 空间的表征) (X,τ)是 T_0 型拓扑空间，当且仅当 X 中的任意两个不同的单点集有不同的闭包，亦即，若 $x,y\in X,x\neq y$，则$\overline{\{x\}}\neq\overline{\{y\}}$.

证 必要性 设 X 是 T_0 空间，$\forall x,y\in X,x\neq y$，例如，$\exists G\in\tau$，s.t. $y\notin G$，故 $G\cap\{y\}=\varnothing$，从而 $x\notin\overline{\{y\}}$，这蕴含$\overline{\{x\}}\neq\overline{\{y\}}$.

充分性 设$\forall x,y\in X,x\neq y$ 有$\overline{\{x\}}\neq\overline{\{y\}}$. 故或$\overline{\{x\}}\backslash\overline{\{y\}}\neq\varnothing$，或$\overline{\{y\}}\backslash\overline{\{x\}}\neq\varnothing$. 若$\overline{\{x\}}\backslash\overline{\{y\}}\neq\varnothing$，则 $x\notin\overline{\{y\}}$(否则，若 $x\in\overline{\{y\}}$，即$\{x\}\subset\overline{\{y\}}$，故$\overline{\{x\}}\subset\overline{\{y\}}$，从而$\overline{\{x\}}\backslash\overline{\{y\}}=\varnothing$)，由此知存在 x 的不含 y 的开集 $\complement\overline{\{y\}}$($x\in\complement\overline{\{y\}}$)；若$\overline{\{y\}}\backslash\overline{\{x\}}\neq\varnothing$，则同理知存在 y 的不含 x 的开集. 故 X 是 T_0 空间.

定理 3.4.2（**T_1 空间的表征**） (X,τ)是 T_1 型拓扑空间，当且仅当 X 中的单点集为闭集.

证 必要性 设 X 是 T_1 空间，$\forall y\in X, y\neq x$，$\exists G_y\in\tau, s.t.\ x\notin G_y$，故 $G_y\cap\{x\}=\varnothing$，从而 $y\notin\overline{\{x\}}$，这蕴含$\overline{\{x\}}=\{x\}$，亦即单点集是闭集.

充分性 设空间 X 中单点集是闭集，则$\forall x,y\in X, x\neq y$，知 $\complement\{x\}$，$\complement\{y\}$为开集，且 $y\in\complement\{x\}$为包含 y 的开集，$x\in\complement\{y\}$为包含 x 的开集；从而 X 是 T_1 空间.

定理 3.4.3（**T_2 空间中极限的惟一性定理**） 设(X,τ)是 T_2 型拓扑空间，则其中任一收敛序列只有一个极限.

证 设(X,τ)为 T_2 型拓扑空间，取定一个序列$\{x_n\}\subset X$，有

$$\lim_{n\to+\infty}x_n=x \quad 与 \quad \lim_{n\to+\infty}x_n=y,$$

且 $x\neq y$. 因为 X 是 T_2 型的，故存在 x 的开邻域 U，y 的开邻域 V，使得 $U\cap V=\varnothing$. 再由极限的定义，对于 $\lim\limits_{n\to+\infty}x_n=x$，存在 $N_x>0$，使得当 $n>N_x$，有 $x_n\in U$；另一方面，对于 $\lim\limits_{n\to+\infty}x_n=y$，存在 $N_y>0$，使得当 $n>N_y$，有 $x_n\in V$. 取 $N=\max\{N_x,N_y\}$，当 $n>N$ 时，$x_n\in U$ 与 $x_n\in V$. 于是，$U\cap V\neq\varnothing$，这与 $U\cap V=\varnothing$矛盾. 从而，只有 $x=y$，亦即，极限若存在，则惟一.

定理 3.4.4（**正规空间的表征**） (X,τ)是正规空间，当且仅当对于 X 的任意闭集 $F\subset X$ 与包含 F 的开集 $G\supset F$，存在一个开集 $O\in\tau$，使得

$$F\subset O\subset\bar{O}\subset G.$$

证 必要性 设 X 是正规空间，对于任意闭集 $F\subset X$，与包含它的任意开集 $G\supset F$，则 $\complement G$ 为闭集，且 $\complement G\subset\complement F$ 给出 $F\cap\complement G=\varnothing$；现在对于两个互不相交的闭集 F 与 $\complement G$，由 X 为正规空间，存在包含 F 的开集 O，$F\subset O$，与包含 $\complement G$ 的开集 W，$\complement G\subset W$，且 $O\cap W=\varnothing$，从而 $O\subset\complement W$，因此有

$$F\subset O\subset\bar{O}\subset\overline{\complement W}=\complement W\subset\complement(\complement G)=G,$$

这蕴含 $F\subset O\subset\bar{O}\subset G$.

充分性 对于 X 的任意两互不相交的闭集 $F_1\subset X$，$F_2\subset X$，$F_1\cap F_2=\varnothing$，知 $\complement F_2\supset F_1$ 是包含 F_1 的一个开集，于是，对于 F_1 与 $\complement F_2$，由假设条件，存在开集 O，使得 $F_1\subset O\subset\bar{O}\subset\complement F_2$. 这样，对于两互不相交的闭集 F_1，F_2，分别存在包含它们的开集 O 与 $\complement\bar{O}$（$\bar{O}\subset\complement F_2$ 给出 $F_2\subset\complement\bar{O}$），且 $O\cap\complement\bar{O}=\varnothing$，故 X 是正规空间.

下面两个是正规空间中的著名定理.

定理 3.4.5（**Urysohn 引理**） (X,τ)是正规空间，当且仅当对于 X 中任意两个互不相交的闭集 A，B，存在 X 上的实值连续函数 $f:X\to\mathbb{R}$，满足

$$f(x)\big|_{x\in A}=0,\quad f(x)\big|_{x\in B}=1, \tag{3.4.1}$$

且 $0\leqslant f(x)\leqslant 1$，$\forall x\in X$（见图 3.4.4）.

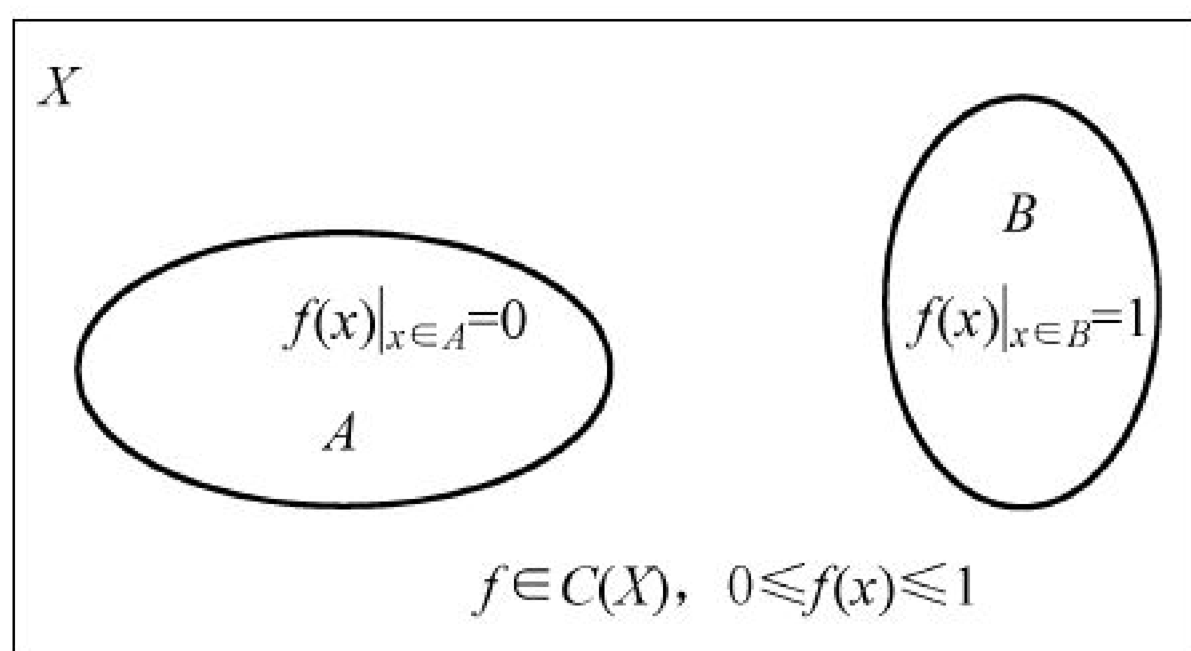


图 3.4.4 Urysohn 引理示意图

证 充分性 当充分性满足时，我们证明(X,τ)是正规空间.

对于两个互不相交的闭集 A,B，若存在实值连续函数 $f:X\to\mathbb{R}$，满足 $f(x)|_{x\in A}=0$，$f(x)|_{x\in B}=1$ 与 $0\leqslant f(x)\leqslant 1$. 我们构造两个集合 $U=\left\{x\in X:f(x)<\frac{1}{2}\right\}$ 与 $V=\left\{x\in X:f(x)>\frac{1}{2}\right\}$，由于 $f:X\to\mathbb{R}$ 的连续性，两者都是开集，并且 $U\cap V=\varnothing$；显然，$A\subset U$，$B\subset V$，从而(X,τ)是正规空间.

必要性 当(X,τ)为正规空间时，对于两个互不相交的闭集 A,B，构造函数 $f:X\to\mathbb{R}$，使其满足所需的条件.

(X,τ) 为正规空间 $\Rightarrow A\cap B=\varnothing$ 蕴含 $A\subset \complement B=X\backslash B$

$\Rightarrow$ 依定理 3.4.4，对于 $A\subset\complement B$，存在开集 $U_{\frac{1}{2}}\subset X$，使得 $A\subset U_{\frac{1}{2}}\subset\overline{U_{\frac{1}{2}}}\subset\complement B$

$\Rightarrow$ 再依定理 3.4.4，对于 $A\subset U_{\frac{1}{2}}$，$\overline{U_{\frac{1}{2}}}\subset\complement B$，存在开集 $U_{\frac{1}{4}}$ 与 $U_{\frac{3}{4}}$，使得

$$A\subset U_{\frac{1}{4}}\subset\overline{U_{\frac{1}{4}}}\subset U_{\frac{1}{2}}\subset\overline{U_{\frac{1}{2}}}\subset U_{\frac{3}{4}}\subset\overline{U_{\frac{3}{4}}}\subset\complement B;$$

如此继续下去，得到 $A\subset U_{\frac{1}{2^n}}\subset\overline{U_{\frac{1}{2^n}}}\subset\cdots\subset U_{\frac{m-1}{2^n}}\subset\overline{U_{\frac{m-1}{2^n}}}\subset\cdots\subset U_{\frac{2^n-1}{2^n}}\subset\overline{U_{\frac{2^n-1}{2^n}}}\subset\complement B$. 于是，若定义 $U_0=A,U_1=\complement B$，则对任一个 $r\in\mathbb{Q}_0=\left\{\frac{m}{2^n}:m=0,1,\cdots,2^n,n\in\mathbb{Z}^+\right\}$，有一个开集 U_r 与之对应，约定 $r>1$ 时 $U_r=X$，最后得

$$U_0\subset U_{\frac{1}{2^n}}\subset\overline{U_{\frac{1}{2^n}}}\subset\cdots\subset U_{\frac{m-1}{2^n}}\subset\overline{U_{\frac{m-1}{2^n}}}\subset\cdots\subset U_{\frac{2^n-1}{2^n}}\subset\overline{U_{\frac{2^n-1}{2^n}}}\subset U_1\subset U_r=X,$$

它们满足下面三条性质：

(1) $r,s\in\mathbb{Q}_0$，$r<s\Rightarrow\overline{U_r}\subset U_s$；

(2) $\forall r\in\mathbb{Q}_0\Rightarrow A=U_0\subset U_r\subset\overline{U_r}\subset\complement B\Rightarrow B\subset\complement(\overline{U_r})$；

(3) $\forall x\in X\Rightarrow x\in B$ 或 $x\in\complement B\Rightarrow x\in B$ 或 $\exists r_x\in\mathbb{Q}_0$，s. t. $x\in\overline{U}_{r_x}$.

现在定义 X 上的实值函数 $f:X\to\mathbb{R}$，满足

$$f(x)=\begin{cases}\inf\{r\in\mathbb{Q}_0:x\in U_r\}, & x\in X\backslash B,\\ 1, & x\in B.\end{cases}\tag{3.4.2}$$

易见，如此定义的 f 有如下性质.

① $\forall x \in A = U_0 \subset U_r$，有 $\inf\{r : x \in U_r, \forall r \in \mathbb{Q}_0\} = 0 \Rightarrow f(x) = 0, \forall x \in A$；

② $\forall x \in B \Rightarrow f(x) = 1$；

③ $\forall x \in X \Rightarrow 0 \leqslant f(x) \leqslant 1$.

余下只要证明 $f(x)$的连续性.由所构造的 $f: X \to [0,1]$是 X 到$\mathbb{R}$ 的子空间$[0,1] \subset \mathbb{R}$ 的映射，故只要证明，对于子空间$([0,1], \tau_{[0,1]})$的子拓扑 $\tau_{[0,1]}$，对任一开集 $O \in \tau_{[0,1]}$，集合 $f^{-1}(O)$是 X 中的开集.易见，只需对于子拓扑 $\tau_{[0,1]}$的一个拓扑基$\mathbb{B}_{[0,1]}$加以证明就够了.

由例 3.2.3，知$(\mathbb{R}, \tau)$有一个子基 $\mathbb{S} = \{\{(-\infty, a) : a \in \mathbb{R}\} \cup \{(b, +\infty) : b \in \mathbb{R}\}\}$，于是，子空间$[0,1]$上有子基$\mathbb{S}_{[0,1]} = \{\{(a,1] : a \in [0,1)\} \cup \{[0,b) : b \in (0,1]\}\}$. 为验证 $f: X \to [0,1]$的连续性，也只要证明$\forall O \in \mathbb{S}_{[0,1]}$，集合 $f^{-1}(O)$是 X 中的开集就够了.

只需考虑 $a \in [0,1)$与 $b \in (0,1]$时，逆像集 $f^{-1}((a,1])$与 $f^{-1}([0,b))$的情况.

当 $a \in [0,1)$时，$f^{-1}((a,1]) \Leftrightarrow f(x) \in (a,1] \Leftrightarrow a < f(x) \leqslant 1 \Leftrightarrow a < \inf\{r \in \mathbb{Q}_0 : x \in U_r\}$；或 $x \in B \Leftrightarrow \exists r \in \mathbb{Q}_0$，s. t. $r > a$，而 $x \notin \overline{U_r} (x \in \mathfrak{C}(\overline{U_r}))$；或 $x \in B \Rightarrow x \in f^{-1}((a,1]) = \left\{\bigcup_{r>a, r \in \mathbb{Q}_0} \mathfrak{C}(\overline{U_r})\right\} \cup B = \left\{\bigcup_{r>a, r \in \mathbb{Q}_0} \mathfrak{C}(\overline{U_r})\right\}$（因 $B \subset \mathfrak{C}(\overline{U_r})$）$\Rightarrow f^{-1}((a,1])$是一族开集$\{\mathfrak{C}(\overline{U_r}) : r \in \mathbb{Q}_0, r > a\}$的并集，故为$(X, \tau)$中的开集.

当 $b \in (0,1]$时，$x \in f^{-1}([0,b)) \Leftrightarrow f(x) \in [0,b) \Leftrightarrow 0 \leqslant f(x) < b \Leftrightarrow \inf\{r \in \mathbb{Q}_0 : x \in U_r\} < b \Leftrightarrow \exists r \in \mathbb{Q}_0$，s. t. $r < b, \forall x \in U_r \Rightarrow x \in f^{-1}([0,b)) = \bigcup_{r<b, r \in \mathbb{Q}_0} U_r \Rightarrow f^{-1}([0,b))$是一族开集$\{U_r : r \in \mathbb{Q}_0, r < b\}$的并集，故为$(X, \tau)$中的开集.定理得证.

此定理的意义在于，把拓扑空间的型与函数(用构造性方法)联系起来，也就是把形体的几何性质用代数方法表示出来(几何方法与代数方法的结合).

定理 3.4.6(Tietze 扩张定理) 设(X, τ)是正规空间，当且仅当对于 X 中任意闭集 A 与定义在 A 上的任一实值连续函数 $f: A \to \mathbb{R}$，存在 X 上的连续函数 $F: X \to \mathbb{R}$，满足 $F(x)|_{x \in A} = f(x)$.

证明较为复杂，读者可参看[12]，[14].

3.4.2 拓扑空间的连通性

连通性，顾名思义，有很直观的几何意义.

例如，集合$[1,2) \cup (2,3]$被分为两个部分，被"隔离"开来；而集合

$$[1,2) \cup [2,3] = [1,3]$$

却是"通"的."连通"反映了"连成一片"的几何直观.

定义 3.4.3(子集的隔离性) 设(X, τ)是拓扑空间. A, B 是 X 的两个子集，若

$$(A \cap \overline{B}) \cup (\overline{A} \cap B) = \varnothing, \tag{3.4.3}$$

则称 A, B **互相隔离(separated each other)**.

注 条件$(A\cap\bar{B})\cup(\bar{A}\cap B)=\varnothing$等价于$A\cap\bar{B}=\varnothing$与$\bar{A}\cap B=\varnothing$同时成立(如图3.4.5所示),可以理解为集合$A,B$互不相交,而且$A$不包含$B$的聚点($A\cap\bar{B}=\varnothing$),$B$也不包含$A$的聚点($\bar{A}\cap B=\varnothing$).

于是,上面的例子是说,$[1,2),(2,3]$互相隔离,而$[1,2),[2,3]$不是隔离的.

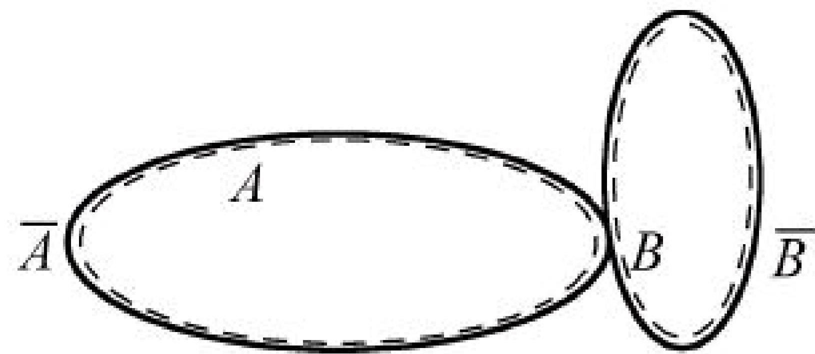


图3.4.5 $A\cap\bar{B}=\varnothing,\bar{A}\cap B=\varnothing$

定义3.4.4(空间的连通性) 设(X,τ)为拓扑空间,若存在非空子集$A,B\subset X$,使得A,B彼此隔离,而且$A\cup B=X$,则称拓扑空间X是**不连通的(disconnected)**;否则,称X是**连通的(connected)**.

拓扑空间(X,τ)的非空子集$A\subset X$称为**不连通子集**,若A作为一个子空间是一个不连通空间.

下面是不连通性的等价性定理.

定理3.4.7 设(X,τ)是拓扑空间,则下述条件等价:

(1) X是不连通空间;

(2) X中存在两个非空闭集$A,B\subset X$,使得$A\cap B=\varnothing,A\cup B=X$;

(3) X中存在两个非空开集$G,O\subset X$,使得$G\cap O=\varnothing,G\cup O=X$;

(4) X中存在既开又闭的非空真子集.

因此,连通空间又可定义为:**若拓扑空间(X,τ)中既开又闭的集合只有X与$\varnothing$,则拓扑空间(X,τ)是连通的.**

证(1)⇒(2) 设X是连通空间,由定义,存在X中两个非空隔离子集$A,B\subset X$,$A\cap\bar{B}=\varnothing,\bar{A}\cap B=\varnothing$, s. t. $A\cup B=X\Rightarrow \complement(A\cup B)=\complement X=\varnothing\Rightarrow A\cap B=\varnothing$. 又由

$$\bar{B}=\bar{B}\cap X=\bar{B}\cap(A\cup B)=(\bar{B}\cap A)\cup(\bar{B}\cap B)=(\bar{B}\cap A)\cup B=\varnothing\cup B=B,$$

$$\bar{A}=\bar{A}\cap X=\bar{A}\cap(A\cup B)=(\bar{A}\cap A)\cup(\bar{A}\cap B)=A\cup(\bar{A}\cap B)=A\cup\varnothing=A,$$

故A,B为闭集,且$A\cap B=\varnothing,A\cup B=X$,因此(2)成立.

(2)⇒(3) 设X中存在两个非空隔离闭集$A,B\subset X$,使$A\cup B=X,A\cap B=\varnothing\Rightarrow \complement A\cap\complement B=\varnothing,\complement A\cup\complement B=X\Rightarrow$存在两非空开集$G=\complement A,O=\complement B$,且$G\cap O=\varnothing,G\cup O=X\Rightarrow$(3)成立.

(3)⇒(4) 若(3)成立,则满足(3)中条件的开集O也是闭集(由$G\cup O=X$知$O=\complement G$为闭集),因此X中存在既开又闭的子集,故(4)成立.

(4)⇒(1) 若X中存在既开又闭的非空真子集A,令$B=\complement A$,则A,B均为X中的既开又闭的非空真子集. 且$A\cup B=X,A\cap B=\varnothing$. 因为两个交为空集的非空闭集一定是隔离的,即$\bar{A}\cap B=A\cap B=\varnothing,A\cap\bar{B}=A\cap B=\varnothing$,因此(1)成立. 定理得证.

例3.4.2 一维欧氏空间$(\mathbb{R},\tau)$是连通空间,有理数集$\mathbb{Q}$是$\mathbb{R}$中的不连通子集.

事实上,由反证法,设欧氏空间$(\mathbb{R},\tau)$不是连通空间,则由定理3.4.7(2),在$\mathbb{R}$中存在两个非空闭集A,B,使得$A\cap B=\varnothing,A\cup B=\mathbb{R}$. 任取$a\in A,b\in B$,不失一般性,可设$a<$

b，$\forall a,b\in\mathbb{R}$. 于是，令$\tilde{A}=A\cap[a,b]$，$\tilde{B}=B\cap[a,b]$，则$\tilde{A}$，$\tilde{B}$是$\mathbb{R}$中两个非空闭集，分别包含a与b，且$\tilde{A}\cap\tilde{B}=\varnothing$，$\tilde{A}\cup\tilde{B}=[a,b]$. 这说明$[a,b]$是一个不连通子集. 但这是不可能的，因为$[a,b]$是连通子集，因而也是连通子空间. 于是，$(\mathbb{R},\tau)$是连通空间.

$[a,b]$的连通性证明如下：

$\tilde{A}=A\cap[a,b]$是一个有上界的集合，$b\in\tilde{B}$就是$\tilde{A}$的一个上界，因此有上确界$\tilde{b}\equiv\sup\tilde{A}<b$（等号不能成立，因为有界闭集的上确界属于它自身，故$\tilde{b}\in\tilde{A}$. 若$\tilde{b}=b$，就有$b\in\tilde{A}\cap\tilde{B}$，这与$\tilde{A}\cap\tilde{B}=\varnothing$矛盾），于是，区间$(\tilde{b},b]\subset\tilde{B}$. 但因$\tilde{B}$为闭集，故$\tilde{b}\in\tilde{B}$，这又导致$\tilde{b}\in\tilde{A}\cap\tilde{B}$，也与$\tilde{A}\cap\tilde{B}=\varnothing$矛盾. 故$[a,b]$必是连通子集，从而也必是连通子空间.

有理数集$\mathbb{Q}$是不连通子集，证明如下：

因为任一个无理数$s\in\mathbb{R}\setminus\mathbb{Q}$，集合$(-\infty,s)\cap\mathbb{Q}$是$\mathbb{Q}$中的开集，同时也是$\mathbb{Q}$中的闭集，因$(-\infty,s)\cap\mathbb{Q}=(-\infty,s]\cap\mathbb{Q}$（注意$(-\infty,s]$是$\mathbb{R}$中的闭集，因为它的补集$\mathbb{R}\setminus(-\infty,s]=(s,+\infty)$为$\mathbb{R}$中的开集），所以$\mathbb{Q}$作为$\mathbb{R}$的子空间是不连通子空间.

连通性是连续映射下保持不变的性质.

定理 3.4.8 （1）设(X,τ)是连通拓扑空间，(Y,υ)是拓扑空间，$f:X\to Y$是连续映射，则$f(X)\subset Y$是Y中的连通子集；（2）设$(X_j,\tau_j)(j=1,2,\cdots,m)$是连通拓扑空间，则积空间$X=X_1\times X_2\times\cdots\times X_m$是连通拓扑空间.

证 （1）反证，若$f(X)$是Y的不连通子集，则存在非空隔离开子集$A,B\subset Y$，使得$f(X)=A\cup B$. 于是，逆像集$f^{-1}(A)$，$f^{-1}(B)$非空，且$f^{-1}(A)$，$f^{-1}(B)\in\tau$. 进而，有包含关系

$$
\begin{aligned}
&[(f^{-1}(A))\cap\overline{(f^{-1}(B))}]\cup[\overline{(f^{-1}(A))}\cap(f^{-1}(B))]\\
&\quad\subset[(f^{-1}(A))\cap(f^{-1}(\bar{B}))]\cup[(f^{-1}(\bar{A}))\cap(f^{-1}(B))]\\
&\quad=f^{-1}[(A\cap\bar{B})\cup(\bar{A}\cap B)]=\varnothing,
\end{aligned}
$$

故$f^{-1}(A)$，$f^{-1}(B)$是非空隔离子集；并且$f^{-1}(A)\cup f^{-1}(B)=f^{-1}(A\cup B)=f^{-1}(X)=X$，这说明$X$是不连通空间，与假设矛盾.

（2）是显然的.

定义 3.4.5（点的连通性、连通分支） 设(X,τ)是拓扑空间. 对于任意两点$x,y\in X$，若存在非空连通子集$A\subset X$，使得$x,y\in A$，则称点x,y是连通的.

一点的连通分支 设(X,τ)是拓扑空间. 对于点$x\in X$，所有包含点x的连通集的并集称为x的**连通分支（Component of x）**.

全不连通性 若拓扑空间(X,τ)中每个点$x\in X$的连通分支都是单点集$\{x\}\subset X$，则称(X,τ)是**全不连通空间**，或称**全断拓扑空间（totally disconnected topological space）**.

例如,赋予 $G=[0,+\infty)$ 以二进拓扑,则所得的拓扑空间是全断的.

事实上,$G=\{x=(x_{-s},\cdots,x_{-1},x_0,x_1,x_2,\cdots):x_j\in\{0,1\},j=-s,-s+1,\cdots,s\in\mathbb{N}\}$.定义 G 上的运算加法 $\oplus$ 为按位加;零点 $0\in G$ 的拓扑开基为 $\tau=\{G_k:k\in\mathbb{Z}\}$,其中

$$G_k=\{y\in G:y=(y_k,y_{k+1},\cdots),y_k\neq 0,y_j\in\{0,1\},j\geqslant k+1\},$$

则 (G,τ) 成为一个拓扑空间,其拓扑 τ 使得 G 成为一个全断空间.

显然,拓扑空间中点之间的连通关系是一个等价关系(即满足自反性、对称性、传递性).

定义 3.4.6(道路连通性) 设 (X,τ) 是拓扑空间. 若 $\forall x,y\in X$,存在闭区间 $[0,1]\subset\mathbb{R}$ 到 X 的连续映射 $\sigma:[0,1]\to X$,使得 $\sigma(0)=x,\sigma(1)=y$,则称 X 为**道路连通的**;连续映射 σ 称为 X 的一条**道路**,x,y 分别称为 σ 的**起点**与**终点**.

连通空间、道路连通空间都是非常有用的概念,这里不再作详细讨论,只强调以下三点:

(1) 道路连通的拓扑空间必定是连通空间,但反之却不然;

(2) 道路连通性是拓扑不变性(亦即,在连续映射下,道路连通性不变);

(3) 欧氏空间中开集的连通性等价于道路连通性,也等价于折线连通性.

定义 3.4.7(局部连通性) 设 (X,τ) 是拓扑空间. 对于 $x\in X$,设 $\tau_x=\{G\in\tau:x\in G\}$ 是包含 x 的开集族,若 $\forall G\in\tau_x$,$\exists V\in\tau_x$, s. t. $V\subset G$,且 V 是连通子集,则称 (X,τ) **在点 x 是局部连通的**;若 (X,τ) 在每个点 $x\in X$ 都是局部连通的,则称 (X,τ) 为**局部连通空间(locally connected space)**.

这里要指出以下几点:

(1) 在局部连通空间中可定义一点的局部连通分支;局部连通性是拓扑不变性;

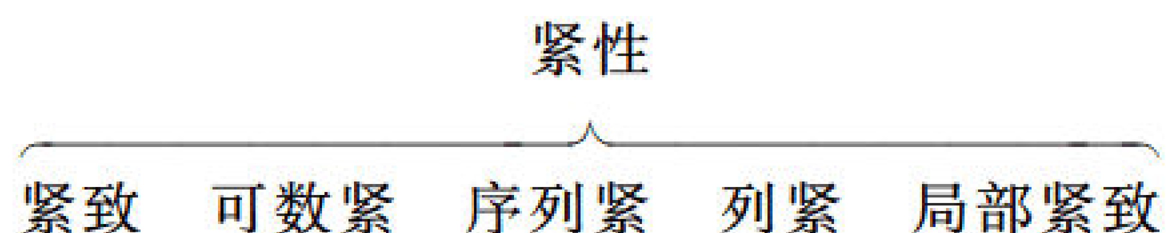
(2) 连通性与局部连通性互不包含;

(3) 可定义局部道路连通空间、局部道路连通分支;

(4) 局部道路连通性在连续映射下不变.

3.4.3 拓扑空间的紧性

紧性在数学研究中有着十分重要的地位. 常用的紧性有以下几种:



本小节重点讨论紧致性与局部紧致性.

1. 紧致性

微积分学中的 Borel 有限覆盖定理可叙述为:"对于任意闭区间 $[\alpha,\beta]$,$-\infty<\alpha<\beta<+\infty$,若 $\mathfrak{D}$ 是 $[\alpha,\beta]$ 的开覆盖,亦即,若 $\mathfrak{D}=\{(a_t,b_t)_{t\in\Lambda}:-\infty<a_t<b_t<+\infty\}$,且 $[\alpha,\beta]\subset$

$\bigcup\limits_{t\in\Lambda}(a_t,b_t)$,则必定存在有限多个 $t_1,\cdots,t_n\in\Lambda$,使得$[\alpha,\beta]\subset\bigcup\limits_{j=1}^{n}(a_{t_j},b_{t_j})$."这是一个把"无限过程化为有限过程"的重要数学概念与数学思维方法,用它可以解决许多抽象的数学问题.

1) 紧致性定义与判别

定义 3.4.8(紧致性) 设(X,τ)为拓扑空间,称 τ 的开子集族 $\Lambda\subset\tau$ 为 X 的开覆盖,若 $X=\bigcup\limits_{A\in\Lambda}A$. 称拓扑空间$(X,\tau)$为**紧致空间(compact space)**,若 X 的任一开覆盖 Λ 都存在有限子覆盖;亦即,对于 X 的任一个开覆盖 Λ,$X=\bigcup\limits_{A\in\Lambda}A$,都存在有限多个开集 $O_1,O_2,\cdots,O_n\in\Lambda$,使得 $X=\bigcup\limits_{j=1}^{n}O_j$. 紧致空间也简称为**紧空间**.

若子集 $A\subset X$ 作为 X 的子空间是紧致空间,则称 A 为 X 中的**紧致子集**,简称**紧集(compact set)**.

等价地,若子集 $A\subset X$ 的任意开覆盖都存在有限的子覆盖,则称子集 A 为紧致子集. 下面的定理证明了这个等价性.

定理 3.4.9 设(X,τ)为拓扑空间,A 为(X,τ)中的紧致子集,当且仅当 A 在(X,τ)中的任何开覆盖 $\mathfrak{U}=\{U\in\tau\}$,必存在有限子覆盖 $\widetilde{\mathfrak{U}}$.

证 必要性 设 $\mathfrak{U}=\{U\in\tau\}$为 A 在(X,τ)中的任一开覆盖 $\Rightarrow$ 子集族 $\mathfrak{P}=\{U\cap A:U\in\mathfrak{U}\}$是 A 在子拓扑空间(A,τ_A)中的开覆盖 $\Rightarrow$ 由必要性假设,A 为紧致子集,故存在$\{U_j\cap A:U_j\in\mathfrak{U}\}_{j=1}^{n}$, s. t. $\bigcup\limits_{j=1}^{n}(U_j\cap A)$是 A 在子拓扑空间(A,τ_A)中的有限开覆盖 $\Rightarrow$ $\widetilde{\mathfrak{U}}=\{U_j\}_{j=1}^{n}$是 A 在拓扑空间(X,τ)中的开覆盖$\mathfrak{U}$的有限子覆盖,$\bigcup\limits_{j=1}^{n}U_j\supset\bigcup\limits_{j=1}^{n}(U_j\cap A)\supset A$ $\Rightarrow$ 必要性得证.

充分性 为证子集 A 的紧致性,对 A 在子拓扑空间(A,τ_A)中的任一开覆盖 $\mathfrak{U}_A=\{V:V\in\tau_A\}$,因其中 $\forall V\in\tau_A$, $\exists U\in\tau$, s. t. $V=U\cap A$,取所有这样的 $U\in\tau$,令 $\mathfrak{U}=\{U\in\tau:U\cap A=V\in\mathfrak{U}_A\}$ $\Rightarrow$ $\mathfrak{U}$是 A 在拓扑空间(X,τ)中的一个开覆盖 $A\subset\bigcup\limits_{V\in\mathfrak{U}_A}V=\bigcup\limits_{U\in\mathfrak{U}}(U\cap A)$ $\Rightarrow$ 由充分性假设,必存在$\mathfrak{U}$的有限子覆盖 $\widetilde{\mathfrak{U}}=\{U_j:U_j\in\mathfrak{U}\}_{j=1}^{n}$, s. t. $A\subset\bigcup\limits_{j=1}^{n}U_j\Rightarrow A\subset\left(\bigcup\limits_{j=1}^{n}U_j\right)\cap A=\bigcup\limits_{j=1}^{n}(U_j\cap A)$;令 $U_j\cap A=V_j$, $1\leqslant j\leqslant n$ $\Rightarrow$ A 在子拓扑空间(A,τ_A)中的任一开覆盖 $\mathfrak{U}_A=\{V:V\in\tau_A\}$,存在有限子覆盖 $\widetilde{\mathfrak{U}}_A=\{V_j\}_{j=1}^{n}$, $A\subset\bigcup\limits_{j=1}^{n}V_j$ $\Rightarrow$ 充分性得证.

例 3.4.3 $\mathbb{R}$ 中的闭区间$[0,1]$是紧致子集. 但$\mathbb{R}$ 自身却不是紧致空间;$\mathbb{R}^n$ 中的闭方体$[a_1,b_1;\cdots;a_n,b_n]$$(-\infty<a_i<b_i<+\infty)$是紧致子集. 但$\mathbb{R}^n$ 本身却不是紧致空间.

若 X 是非空有限集，τ 是 X 上的拓扑，则(X,τ)必定是紧致空间.

我们只证明$\mathbb{R}$中的闭区间$[0,1]$是紧致子集.

取闭区间$[0,1]$的一个开覆盖$\mathfrak{U}=\{V=U\cap[0,1]:U\in\mathfrak{B}$，这里$(\mathbb{R},\tau)$为$\mathbb{R}$中的通常拓扑，$\mathfrak{B}$为 τ 的拓扑基 $\mathfrak{B}=\{(a,b):-\infty<a<b<+\infty\}$，则$\mathfrak{U}=\{(a,b)\cap[0,1]:(a,b)\in\mathfrak{B}$；令

$$\mathfrak{P}=\{x\in[0,1]:\exists\mathfrak{U}\text{的有限子集族覆盖}[0,x]\}, \tag{3.4.4}$$

下面证明：(1) $\mathfrak{P}\neq\varnothing$；

(2) $\mathfrak{P}$是$[0,1]$中的开集；

(3) $\mathfrak{P}$是$[0,1]$中的闭集.

(1) 显然成立. 为证(2)，由定义，$\forall x\in\mathfrak{P}\ \exists\{V_1,V_2,\cdots,V_n\}\subset\mathfrak{U}$ s. t. $[0,x]\subset\bigcup\limits_{j=1}^{n}V_j$. 分两种情况讨论.

① 若 $x=1\in\mathfrak{P}\Rightarrow$ 因$[0,x]=[0,1]$，而$[0,1]$在子拓扑空间$([0,1],\tau')$中是开集，因此 $x=1$ 是内点；同理，若 $x=0\in\mathfrak{P}$则 0 是内点.

② 若 $x\in\mathfrak{P}$且 $0<x<1\Rightarrow$ 首先由 $0<x<1$，因 $\mathfrak{U}$是$[0,1]$的开覆盖，故 $\exists V_0$，s. t. $x\in V_0\in\mathfrak{U}\Rightarrow$ 因 $V_0=U_0\cap[0,1]$是$[0,1]$中的开集，故 x 是 V_0 的内点 $\Rightarrow\exists\varepsilon_1>0$，s. t. $[x,x+\varepsilon_1)\subset(x-\varepsilon_1,x+\varepsilon_1)\cap[0,1]\subset V_0$；其次由 $x\in\mathfrak{P}\ \exists V_1,V_2,\cdots,V_n\in\mathfrak{U}$ s. t. $[0,x]\subset\bigcup\limits_{j=1}^{n}V_j\Rightarrow$ 取 $\varepsilon(0<\varepsilon<\varepsilon_1)$，使$[0,x+\varepsilon]=[0,x]\cup[x,x+\varepsilon_1)\subset\left(\bigcup\limits_{j=1}^{n}V_j\right)\cup V_0\Rightarrow x+\varepsilon\in\mathfrak{P}$且$[0,x+\varepsilon)\subset\mathfrak{P}$是 $\mathfrak{P}$中包含 x 的开集，故 x 是 $\mathfrak{P}$的内点 $\Rightarrow$ 结合①和②，则 $\mathfrak{P}$是开集.

为证(3)，$\mathfrak{P}$是$[0,1]$中的闭集，我们证明 $\complement_{[0,1]}\mathfrak{P}=[0,1]\backslash\mathfrak{P}$为开集.

设 $x\in\complement_{[0,1]}\mathfrak{P}\Rightarrow[x,1]\subset\complement_{[0,1]}\mathfrak{P}\Rightarrow x>0$(因 $0\in\mathfrak{P}\Rightarrow\exists\widetilde{V}\in\mathfrak{U}$ s. t. $x\in\widetilde{V}$($\mathfrak{U}$是$[0,1]$的开覆盖) $\Rightarrow\widetilde{V}$是$[0,1]$中的开集，$\exists\varepsilon>0$，s. t. $(x-\varepsilon,x]\subset(x-\varepsilon,x+\varepsilon)\cap[0,1]\subset\widetilde{V}\Rightarrow(x-\varepsilon,x]\cap\mathfrak{P}=\varnothing$(否则，若$(x-\varepsilon,x]\cap\mathfrak{P}\neq\varnothing\Rightarrow\exists z\in(x-\varepsilon,x]\cap\mathfrak{P}\Rightarrow z\in\mathfrak{P}\Rightarrow\exists$有限子集族 $\mathcal{U}\subset\mathfrak{U}$ s. t. $\mathcal{U}$覆盖$[0,z]\Rightarrow$ 开覆盖$\{\widetilde{V}\}\cup\mathcal{U}$覆盖$[0,z]\cup(x-\varepsilon,x]=[0,x]\Rightarrow x\in\mathfrak{P}$这与 $x\in\complement_{[0,1]}\mathfrak{P}$矛盾) $\Rightarrow(x-\varepsilon,x]\subset\complement_{[0,1]}\mathfrak{P}\Rightarrow(x-\varepsilon,1]=(x-\varepsilon,x]\cup[x,1]\subset\complement_{[0,1]}\mathfrak{P}\Rightarrow x$ 是 $\complement_{[0,1]}\mathfrak{P}$的内点 $\Rightarrow\complement_{[0,1]}\mathfrak{P}$是开集 $\Rightarrow\mathfrak{P}$是闭集.

综上，$\mathfrak{P}$是$[0,1]$中既开又闭的子集，且 $\mathfrak{P}\neq\varnothing$. 然而，在例 3.4.2 中已经证明了$[0,1]$是连通集，故 $\mathfrak{P}$不可能是既开又闭的，除非有 $\mathfrak{P}=[0,1]$. 另一方面，$1\in\mathfrak{P}$表示$\exists\mathfrak{U}$的有限子集族 $\mathcal{U}$，它覆盖$[0,1]$，于是$[0,1]$是一个紧致子集.

例 3.4.4 设拓扑空间(X,τ)的拓扑 τ 是①最粗拓扑、②余有限拓扑(X 为无限集)，则(X,τ)是紧致空间.

事实上，① 最粗拓扑 $\tau=\{X,\varnothing\}$显然使得(X,τ)是紧致空间.

② X 为无限集，余有限拓扑 $\tau=\{G\subset X:\complement G$ 为有限子集$\}$，对 X 的任一个开覆盖 Λ，$X=$

$\bigcup_{O\in\Lambda} O$，取 $O\in\Lambda, O\neq\varnothing$，则 $\complement O$ 为 X 的有限集，$\complement O=\{a_1,a_2,\cdots,a_n\}$. 取 $O_j\in\Lambda, a_j\in O_j, j=1,2,\cdots,n$，则有限集族 $\{O,O_1,\cdots,O_n\}\subset\Lambda$ 是 Λ 的有限子集族，且是 X 的有限开覆盖：$X\subset O\cup O_1\cup\cdots\cup O_n$. 故余有限拓扑使得 (X,τ) 是紧致空间.

为判断拓扑空间 (X,τ) 的紧致性，可以用它的一个拓扑基来进行.

定理 3.4.10　设 $\mathfrak{B}$ 是拓扑空间 (X,τ) 的一个拓扑基，若 X 的任何一个由 $\mathfrak{B}$ 中集合所构成的开覆盖都有有限子覆盖，则 (X,τ) 是紧致拓扑空间.

证　设 $\mathfrak{U}=\{U\subset X: U\in\mathfrak{B}\}$ 是 X 的所述的开覆盖 $\Rightarrow \forall U\in\mathfrak{U}, \exists\, \mathfrak{B}$ 的子集族 $\mathfrak{B}_U$, s. t. $U=\bigcup_{B\in\mathfrak{B}_U} B \Rightarrow$ 令 $\widetilde{\mathfrak{U}}=\bigcup_{U\in\mathfrak{U}}\mathfrak{B}_U$，则 $\bigcup_{B\in\widetilde{\mathfrak{U}}} B=\bigcup_{B\in\bigcup_{U\in\mathfrak{U}}\mathfrak{B}_U} B=\bigcup_{U\in\mathfrak{U}}\left(\bigcup_{B\in\mathfrak{B}_U} B\right)=\bigcup_{U\in\mathfrak{U}} U=X \Rightarrow \widetilde{\mathfrak{U}}$ 是由 $\mathfrak{B}$ 中的元构成的 X 的开覆盖，据假设，存在 X 的有限开覆盖 $\{B_j\in\widetilde{\mathfrak{U}}: j=1,2,\cdots,n\}$, s. t. $\bigcup_{j=1}^{n} B_j=X \Rightarrow$ 由 $B_j\in\widetilde{\mathfrak{U}}(j=1,2,\cdots,n)$，$\exists\, U_j\in\mathfrak{U}$ s. t. $B_j\in\mathfrak{B}_{U_j}(j=1,2,\cdots,n) \Rightarrow B_j\subset U_j\in\mathfrak{U}\Rightarrow$ 对于 $\mathfrak{U}$ 的有限子集族 $\{U_j: j=1,2,\cdots,n\}\subset\mathfrak{U}$ 有 $\bigcup_{j=1}^{n} U_j\supset\bigcup_{j=1}^{n} B_j=X \Rightarrow$ 定理得证.

2）紧致性与连续性的关系

定理 3.4.11　(1) 设 (X,τ)，(Y,υ) 是拓扑空间，$f: X\to Y$ 是连续映射，若 $A\subset X$ 是 X 中的紧致子集，则 $f(A)\subset Y$ 是 Y 中的紧致子集；亦即，连续映射保持紧致性；

(2) 设 $(X_j,\tau_j)(j=1,2,\cdots,m)$ 是紧致空间，则积空间 $X=X_1\times X_2\times\cdots\times X_m$ 是紧致空间.

证　(1) 设 $A\subset X$ 是 X 中的紧致子集. 对于像集 $f(A)\subset Y$，设 $\mathfrak{U}=\{U\subset Y: U\in\upsilon\}$ 是 $f(A)$ 的任一个开覆盖，$f(A)\subset\bigcup_{U\in\mathfrak{U}} U$. 由 f 的连续性知，$\forall U\in\mathfrak{U}, f^{-1}(U)\subset X$ 是 X 中的开子集，并且

$$A\subset f^{-1}(f(A))\subset f^{-1}\left(\bigcup_{U\in\mathfrak{U}} U\right)=\bigcup_{U\in\mathfrak{U}} f^{-1}(U). \tag{3.4.5}$$

于是，$\bigcup_{U\in\mathfrak{U}} f^{-1}(U)$ 是 A 的一个开覆盖. 由 A 的紧致性，存在 $\mathfrak{U}$ 中有限多个开集 $U_1,U_2,\cdots,U_n\in\mathfrak{U}$ 使得 $f^{-1}(U_1), f^{-1}(U_2),\cdots, f^{-1}(U_n)$ 形成 A 的开覆盖 $A\subset\bigcup_{j=1}^{n} f^{-1}(U_j)=f^{-1}\left(\bigcup_{j=1}^{n} U_j\right)$，故

$$f(A)\subset f\left(f^{-1}\left(\bigcup_{j=1}^{n} U_j\right)\right)\subset\bigcup_{j=1}^{n} U_j,$$

从而 $f(A)$ 是 Y 中的紧致子集.

(2) 不失一般性，对 $j=1,2$ 进行证明. 设 (X,τ)，(Y,υ) 为紧致空间，由积空间的定义，知集合 $\mathfrak{B}=\{U\times V:U\in\tau,V\in\upsilon\}$ 是积空间 $(X\times Y,\tau\times\upsilon)$ 的一个拓扑基，据定理 3.4.10，取 $\mathfrak{B}$ 中的元所构成的 $X\times Y$ 构成的开覆盖 $\mathfrak{U}=\{U\times V\subset X\times Y:U\times V\in\mathfrak{B}$. 于是 $\forall x\in X$，子空间 $\{x\}\times Y\xleftrightarrow{\text{同胚}}Y\Rightarrow\{x\}\times Y$ 是紧致子空间蕴含 $\{x\}\times Y$ 是 $X\times Y$ 的紧致子集 $\Rightarrow$ 据假设 $\mathfrak{U}$ 是 $X\times Y$ 的开覆盖，显然也是紧致子集 $\{x\}\times Y$ 的开覆盖，故 $\exists\mathfrak{U}$ 的有限子覆盖，记为 $\widetilde{\mathfrak{U}}=\{U_x\times V_{y_1},\cdots,U_x\times V_{y_n}\subset\{x\}\times Y:U_x\times V_{y_j}\in\mathfrak{U}$，且 $U_x\times V_{y_j}$ 与 $\{x\}\times Y$ 的交非空(否则可剔除)s.t. $(U_x\times V_{y_1})\cup\cdots\cup(U_x\times V_{y_n})=U_x\times(V_{y_1}\cup\cdots\cup V_{y_n})\supset U_x\times Y\Rightarrow$ 因 $\bigcup\limits_{x\in X}U_x=X$ 紧致，故 $\exists U_{x_1},\cdots,U_{x_m}$，且 $U_{x_j}\times V_{y_k}\in\widetilde{\mathfrak{U}}$，s.t. $U_{x_1}\cup\cdots\cup U_{x_m}=X$，且 $(U_{x_1}\cup\cdots\cup U_{x_m})\times Y=X\times Y\Rightarrow$ $\{U_{x_j}\times V_{y_k}:U_{x_j}\times V_{y_k}\in\widetilde{\mathfrak{U}},1\leqslant j\leqslant m,1\leqslant k\leqslant n\}\supseteq X\times Y\Rightarrow X\times Y$ 为紧致空间.

3) 紧致集与闭集的关系

定理 3.4.12 设 (X,τ) 是拓扑空间.

(1) 若 X 是紧致空间，则 X 的闭子集 A 是 X 中的紧致子集；

(2) 若 X 是 T_2 型空间，则 X 的紧致子集 A 是 X 中的闭子集；

(3) 若 X 是 T_2 型紧致空间，则 X 中子集 A 是紧致子集，当且仅当 A 是闭子集.

证 对于(1)，设 $A\subset X$ 是紧致空间 X 中的闭子集，并设 $\mathfrak{U}=\{V\subset X:V\in\tau\}$ 是 A 的任一个开覆盖，$A\subset\bigcup\limits_{V\in\mathfrak{U}}V$. 由于 A 是闭子集，故 $\complement A$ 为开集，于是 $\left(\bigcup\limits_{V\in\mathfrak{U}}V\right)\cup\complement A=\left(\bigcup\limits_{V\in\mathfrak{U}}V\right)\cup(X\backslash A)=X$，这里 $\mathfrak{U}=\left\{\left(\bigcup\limits_{V\in\mathfrak{U}}V\right)\cup\complement A\right\}$ 是 X 的开覆盖，由于 X 的紧致性，存在有限开覆盖 $B_j\in\mathfrak{U}(j=1,2,\cdots,n)$，使得 $X=\bigcup\limits_{j=1}^{n}B_j$，$B_j\in\mathfrak{U}$. 显然，$B_j\in\mathfrak{U}$ 中必有一个是 $\complement A$，例如 $B_n=\complement A$，因此 $A\subset\{B_j\}_{j=1}^{n-1}\subset\mathfrak{U}$，所以 A 的任一开覆盖 $\mathfrak{U}$ 都存在有限子覆盖 $\{B_j\}_{j=1}^{n-1}$，故闭子集 A 是紧致的.

为证(2)，设 X 是 T_2 型拓扑空间，$A\subset X$ 为紧致子集，欲证 A 为闭集，分为两步.

① 点 $x\notin A$ 与紧致集 A 的分离性　对于任一点 $x\in X$ 与紧致集 A，若 $x\notin A$，则存在包含 x 的开集 U，$x\in U$，与包含 A 的开集 V，使得 $U\cap V=\varnothing$.

事实上，对于 $x\in X$ 与任意的 $y\in A$，由于空间是 T_2 型的，故存在包含 x 的开邻域 $U_y(x)$，与包含 y 的开邻域 V_y，使得 $U_y(x)\cap V_y=\varnothing$. 显然，集族 $\{V_y:y\in A\}$ 是集 A 的开覆盖，$A\subset\bigcup\limits_{y\in A}V_y$. 据 A 的紧致性，存在有限多个开集 $V_{y_1},V_{y_2},\cdots,V_{y_n}$，使得 $A\subset\bigcup\limits_{j=1}^{n}V_{y_j}\equiv V$. 相应地，有 $U_{y_1}(x),U_{y_2}(x),\cdots,U_{y_n}(x)$. 令 $U\equiv U(x)=\bigcap\limits_{j=1}^{n}U_{y_j}(x)$，$U$ 是包含 $x\in U$ 的开集，

$V=\bigcup_{k=1}^{n}V_{y_k}$ 则是包含 A 的开集，且 $U\cap V_{y_k}=\left(\bigcap_{j=1}^{n}U_{y_j}\right)\cap V_{y_k}=\bigcap_{j=1}^{n}(U_{y_j}\cap V_{y_k})=\varnothing$ 与 $U\cap V=U\cap\bigcup_{k=1}^{n}V_{y_k}=\bigcup_{k=1}^{n}(U\cap V_{y_k})=\varnothing$.

② 紧致集 A 是闭集　对于任一点 $x\in X$，若 $x\notin A$，则 x 不是 A 的聚点，从而 A 是闭集.

事实上，由①，存在 x 的开邻域 U，使得 U 中不含 A 的点，因此 $x\notin A$ 不可能是 A 的聚点. 于是，A 的聚点全部包含在 A 中，所以 A 是闭集.

(3) 由(1)、(2)得到.

归纳地，我们得到：

在紧致空间中，闭子集 $\Rightarrow$ 紧致子集；

在 T_2 型空间中，闭子集 $\Leftarrow$ 紧致子集；

在 T_2 型紧致空间中，闭子集 $\Leftrightarrow$ 紧致子集.

例 3.4.5　拓扑空间(X,τ)中的紧致集但非闭集的例子.

① 若 X 是至少含两个点的有限集，τ 是最粗拓扑 $\tau=\{X,\varnothing\}$——任一点 $x\in X$，子集 $A=X\setminus\{x\}$是 X 中的紧致集(因它是有限集)，非闭集(X 中仅两个闭集 X 与$\varnothing$).

② 设 X 是无限集，$(X,\tau_{余有限})$是余有限拓扑——任一点 $x\in X$，子集 $A=X\setminus\{x\}$是紧致集，但却非闭，因$(X,\tau_{余有限})$是 T_1 空间，单点集$\{x\}$是闭集，故补集 $A=X\setminus\{x\}$为开集.

定理 3.4.13　设(X,τ)是紧致拓扑空间，(Y,υ)是 T_2 型拓扑空间，则

(1) 连续映射 $f:X\to Y$ 是闭映射；

(2) 连续的一一映射 $f:X\to Y$(单射加满射)是同胚映射.

证　(1) 为证 $f:X\to Y$ 为闭映射，任取闭子集 $A\subset X\Rightarrow A$ 为紧致集(由 X 的紧致性与定理 3.4.12(1)) $\Rightarrow f(A)\subset Y$ 为 Y 中的紧致子集(由 f 的连续性与定理 3.4.11(1)) $\Rightarrow f(A)$为 Y 中的闭子集(由 Y 为 T_2 型与定理 3.4.12(3)) $\Rightarrow f$ 为闭映射.

(2) 为证连续的一一映射 $f:X\to Y$ 为同胚映射，记 f 的逆映射为 f^{-1}，任取闭子集 $A\subset X\Rightarrow(f^{-1})^{-1}(A)=f(A)$，$f(A)$为闭集 $\Rightarrow f^{-1}$的逆像集$(f^{-1})^{-1}(A)$是闭集 $\Rightarrow f^{-1}$为连续映射 $\Rightarrow f:X\to Y$ 双方单值、双方连续，故为同胚映射.

4) 紧致性与分离性的关系

定理 3.4.14　设(X,τ)是 T_2 型拓扑空间.

(1) 若 $A\subset X$ 为紧致子集，且 X 中的点 $x\notin A$，则分别存在它们的开邻域 $U\in\tau$ 与 $V\in\tau$，使得 $U\cap V=\varnothing$(点与紧致子集的分离性)；

(2) 若 $A\subset X$ 与 $B\subset X$ 为两个紧致子集，且 $A\cap B=\varnothing$，则存在 A 与 B 的开邻域 $U\in\tau$ 与 $V\in\tau$，使得 $U\cap V=\varnothing$(两紧致子集的分离性).

证　(1) 在定理 3.4.12(2)的证明中已经得到这个结果.

(2) 对于紧致子集 $A\subset X, B\subset X, A\cap B=\varnothing \Rightarrow \forall x\in A$ 与紧致集 B，由(1)，$\exists x$ 的开邻域 $U_x\in\tau, x\in U_x$，与紧致集 B 的开邻域 $V_x\equiv V_x\in\tau, V_x\supset B$，s. t. $U_x\cap V_x=\varnothing \Rightarrow \{U_x: x\in A\}$是 A 的开覆盖，由 A 的紧致性，$\exists$ 有限子覆盖 $A\subset \bigcup_{j=1}^{n} U_{x_j}$；相应地，$V_{x_j}\supset B, 1\leqslant j\leqslant n$，s. t. $U_{x_j}\cap V_{x_k}=\varnothing, 1\leqslant k, j\leqslant n \Rightarrow$ 令 $U\equiv \bigcup_{j=1}^{n} U_{x_j}, V=\bigcap_{k=1}^{n} V_{x_k} \Rightarrow A\subset U, B\subset V$，且 $U\cap V=\varnothing \Rightarrow$ (2)成立.

定理 3.4.15 设(X,τ)是拓扑空间.

(1) 若 X 是 T_2 型紧致空间，则 X 是正则空间；

(2) 若 X 是 T_2 型紧致空间，则 X 是正规空间；

(3) 若 X 是紧致的正则空间，则 X 是正规空间.

证 (1) 为证 X 是正则空间，任取一点 $x\in X$ 与任一闭子集 $A\subset X$，且 $x\notin A \Rightarrow$ 由 X 的紧致性，A 为紧致子集 $\Rightarrow$ 由 X 的 T_2 型与紧致性，据定理 3.4.14(1)，对 $x\notin A$ 与紧致子集 A，存在 x 的开邻域 U 与 A 的开邻域 V，s. t. $U\cap V=\varnothing \Rightarrow (X,\tau)$为正则空间.

(2) 为证 X 是正规空间，任取 X 中的两个互不相交的闭子集 $A, B \Rightarrow$ 由 X 的 T_2 型与紧致性，A, B 为互不相交的紧致子集 $\Rightarrow$ 据定理 3.4.14(2)，存在 A, B 的互不相交的开邻域 U, V，s. t. $U\cap V=\varnothing \Rightarrow X$ 为正规空间.

(3) 设 X 为紧致的正则空间，为证 X 是正规空间，任取闭子集 $F\subset X$ 与包含 F 的开邻域 $G\supset F$，我们证明，存在开集 $O\in\tau$，s. t. $F\subset O\subset\bar{O}\subset G$，于是，据定理 3.4.4，知 X 为正规空间.

事实上，由 X 的紧致性与 $F\subset X$ 是闭子集，故 F 是紧致子集；由 $G\supset F$ 是开邻域，X 是正则空间，因此，$\forall x\in F$ 都存在 x 的开邻域 $V_x\in\tau$，s. t. $x\in V_x\subset\bar{V}_x\subset G$. 于是，开集族$\{V_x: x\in F\}$成为 F 的开覆盖 $F\subset \bigcup_{x\in F} V_x$，据 F 的紧致性，$\exists\{V_1, V_2, \cdots, V_n\}$，s. t. $F\subset \bigcup_{j=1}^{n} V_j$. 令 $V=\bigcup_{j=1}^{n} V_j$，$F\subset V$，它是开邻域，且 $F\subset V\subset\bar{V}=\overline{\bigcup_{j=1}^{n} V_j}=\bigcup_{j=1}^{n}\bar{V}_j\subset G$，从而，定理 3.4.4 给出 X 为正规空间.

例 3.4.6 紧致的正规空间(X,τ)不是正则空间的例子.

令 $X=\{1,2,3\}, \tau=\{X,\varnothing,\{1\},\{2\},\{1,2\}\}$，它是正规的、紧致的，但非正则的.

5) 度量空间中的紧致性

对于度量空间(X,ρ)与欧氏空间$\mathbb{R}^n$，紧致性更有许多重要性质.

定义 3.4.9(有界性) 设(X,ρ)为度量空间，称子集 $A\subset X$ 为**有界集**(**bounded set**)，若存在实数 $M>0$，使得 $\rho(x,y)<M$ 对所有 $x, y\in A$ 都成立. 若 X 本身是有界的，则称 X 为**有界度量空间**.

定理 3.4.16 设(X,ρ)是紧致度量空间,则

(1) X 是有界空间;

(2) 当 $X\neq\varnothing$ 时,连续映射 $f:X\to\mathbb{R}$ 必可达到其最大值与最小值,亦即 $\exists x_{\max}\in X$ 与 $\exists x_{\min}\in X$, s. t.

$$f(x_{\max})=\max_{x\in X}\{f(x)\} \quad 与 \quad f(x_{\min})=\min_{x\in X}\{f(x)\};$$

(3) 若连续映射 $f:X\to Y$,其中(Y,ρ_1)为度量空间,则 f 是一致连续的.

证 (1) 设(X,ρ)是紧致度量空间 $\Rightarrow$ 半径为 1 的球族 $\mathfrak{B}=\{B(x,1):x\in X\}$是 X 的一个开覆盖,由(X,ρ)的紧致性,存在有限子覆盖$\{B(x_1,1),B(x_2,1),\cdots,B(x_n,1)\}$ $\Rightarrow$ 令 $M=\max\{\rho(x_j,x_k):1\leqslant j,k\leqslant n\}+2$ $\Rightarrow$ $\forall x,y\in X$, $\exists j,k(1\leqslant j,k\leqslant n)$, s. t. $x\in B(x_j,1)$, $y\in B(x_k,1)$,且 $\rho(x,y)\leqslant\rho(x,x_j)+\rho(x_j,x_k)+\rho(x_k,y)<1+(M-2)+1=M$ $\Rightarrow$ X 是有界空间.

(2) $X\neq\varnothing$是紧致空间 $\Rightarrow$ $f(X)$为$\mathbb{R}$中的紧致子集 $\Rightarrow$ $f(X)$为$\mathbb{R}$中的有界闭集 $\Rightarrow$ $\inf f(X)\leqslant f(x)\leqslant\sup f(X)$ $\Rightarrow$ 因 $f(X)$是闭集,故 $\inf f(X)=\min\limits_{x\in X}\{f(x)\}$,也有 $\sup f(X)=\max\limits_{x\in X}\{f(x)\}$ $\Rightarrow$ $f(x_{\min})=\min\limits_{x\in X}\{f(x)\}$与 $f(x_{\max})=\max\limits_{x\in X}\{f(x)\}$.

(3) 设 $f:X\to Y$ 是紧致度量空间(X,ρ)到度量空间(Y,ρ_1)的连续映射,欲证 f 一致连续,是指:$\forall\varepsilon>0$, $\exists\delta=\delta(\varepsilon)>0$, s. t. 当 $x',x''\in X$ 且 $\rho(x',x'')<\delta$ 时,有 $\rho_1(f(x'),f(x''))<\varepsilon$.

由 f 的连续性,$\forall x\in X$, $\forall\varepsilon>0$, $\exists\delta=\delta(\varepsilon,x)>0$, s. t. 当 $\rho(x',x)<\delta$ 时,有 $\rho_1(f(x'),f(x))<\frac{\varepsilon}{2}$ $\Rightarrow$ 开球族$\left\{B\left(x,\frac{\delta(\varepsilon,x)}{2}\right):x\in X\right\}$是紧致度量空间$(X,\rho)$的开覆盖 $\Rightarrow$ 存在 X 的有限子覆盖 $\left\{B\left(x_j,\frac{\delta(\varepsilon,x_j)}{2}\right):x_j\in X,j=1,2,\cdots,n\right\}$ $\Rightarrow$ 令 $\delta(\varepsilon)=\min\limits_{1\leqslant j\leqslant n}\{\delta_j\}$ $\Rightarrow$ $\forall x',x''\in X$,当$\rho(x',x'')<\delta$ 时,若 $x'\in B\left(x_{j_0},\frac{\delta(\varepsilon,x_{j_0})}{2}\right)$,则 $x''\in B(x_{j_0},\delta(\varepsilon))$ $\Rightarrow$ $\rho_1(f(x'),f(x''))\leqslant\rho_1(f(x'),f(x_{j_0}))+\rho_1(f(x_{j_0}),f(x''))<\varepsilon$ $\Rightarrow$ $f(x)$在 X 上一致连续. 定理得证.

定理 3.4.17 $\mathbb{R}^n$ 中的子集 A 是紧致子集,当且仅当 A 是有界闭集.

证 这里只证 $n=1$ 情形.

必要性 设 $A\subset\mathbb{R}$ 是紧致子集 $\Rightarrow$ $A\subset\mathbb{R}$ 是有界子集(定理 3.4.16),且是闭集($\mathbb{R}$ 是 T_2 空间,故紧致子集为闭子集).

充分性 设 $A\subset\mathbb{R}$ 是有界闭子集 $\Rightarrow$ 若 $A=\varnothing$,则显然是紧致集;若 $A\neq\varnothing$ $\Rightarrow$ $\exists M>0$, s. t. $\forall x,y\in A$ 有 $\rho(x,y)<M$ $\Rightarrow$ 任取 $x_0\in A$,令 $M_0=M+\rho(0,x_0)$,其中 $0\in\mathbb{R}$ 是原点 $\Rightarrow$ $A\subset[-M_0,M_0]$(由三点不等式得到 $\forall x\in A$,有 $\rho(x,0)\leqslant\rho(x,x_0)+\rho(x_0,0)<M+\rho(x_0,0)=M_0$) $\Rightarrow$ 由$[-M_0,M_0]$为紧致子空间($[-M_0,M_0]\xleftrightarrow{\text{同胚}}[0,1]$)的闭子集是紧致子集,故 A 是紧致集. 定理得证.

2. 局部紧致性

定义 3.4.10(局部紧致性) 设(X,τ)为拓扑空间,若 X 的每个点 $x\in X$ 都存在紧致邻域,亦即,$\forall x\in X$,$\exists X$ 的紧致子集 $W\subset X$,s.t. $x\in\mathring{W}\subset W$($\mathring{W}$ 就是 $x\in X$ 的开邻域),则称拓扑空间(X,τ)为**局部紧致空间(locally compact space)**.

若集合 $A\subset X$ 作为子空间是一个局部紧致空间,则称 A 为 X 中的局部紧致子集,简称**局部紧集(locally compact set)**.

若集 $A\subset X$ 的闭包 $\overline{A}$ 是 X 的紧致子集,则称 A 为**相对紧致集(relative compact set)**.

紧致拓扑空间(X,τ)是局部紧致的,因为 $\forall x\in X$,令 $W=X$,视 W 为 x 的紧致邻域,则 $x\in\mathring{W}=\mathring{X}=X=W$,且 $W=X$ 为紧致集,故(X,τ)是局部紧致空间. 但反之不然.

例 3.4.7 $\mathbb{R}$ 是局部紧致空间,但非紧致空间.

事实上,$\forall x\in\mathbb{R}$,取 $r>0$,则 $\overline{B(x,r)}=\overline{(x-r,x+r)}=[x-r,x+r]$为$\mathbb{R}$ 中的紧致集,且 $x\in B(x,r)=\overset{\circ}{\overline{B(x,r)}}\subset\overline{B(x,r)}$,而 $\overline{B(x,r)}$是 x 的紧致闭包邻域. 但$\mathbb{R}$ 不是紧致空间,因为$\mathbb{R}$ 的开覆盖 $\mathfrak{U}=\{(-n,n):n\in\mathbb{N}\}$不存在$\mathbb{R}$ 的有限子覆盖.

对于有限实数 a,集合$(-\infty,a)$是$\mathbb{R}$ 中的局部紧致子集.

局部紧致空间具有如下性质.

定理 3.4.18 设(X,τ)是局部紧致拓扑空间,则其中的闭子集 $F\subset X$ 是局部紧致子集.

T_2 型局部紧致空间具有如下性质.

定理 3.4.19 设(X,τ)是 T_2 型局部紧致拓扑空间,若 $U\in\tau$ 为开集,则对任一点 $x\in U$,存在 x 的开邻域 V,使得 $x\in V\subset\overline{V}\subset U$,如图 3.4.6 所示. 由此,$T_2$ 型局部紧致拓扑空间是局部紧致的正则空间.

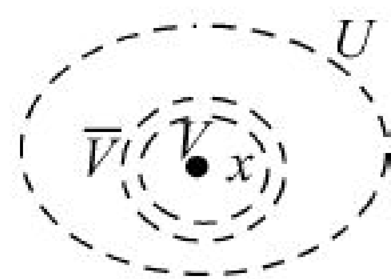


图 3.4.6 分离性

证 X 为局部紧致的 $\Rightarrow x\in U,U\in\tau$,$\exists$紧致子集 $F\subset X$,s.t. $x\in\mathring{F}$,且因 X 为 T_2 型,故 F 为闭集 $\Rightarrow(F,\tau_F)$作为子空间,是 T_2 型紧致空间,故为正则空间(定理 3.4.15) $\Rightarrow x\in U\cap\mathring{F}$ 是 x 在空间(X,τ)中的开邻域,故也是子空间(F,τ_F)中的开邻域 $\Rightarrow\exists x$ 在子空间(F,τ_F)中的开邻域 V,$x\in V$,s.t. V 在子空间(F,τ_F)中的闭包 $\overline{V}_F$满足 $x\in V\subset\overline{V}_F\subset W=U\cap\mathring{F}\Rightarrow\overline{V}_F=\overline{F\cap V}=\overline{F}\cap\overline{V}=F\cap\overline{V}=\overline{V}$(因 F 闭),即 V 在子空间(F,τ_F)中的闭包与在空间(X,τ)中的闭包相等 $\Rightarrow\overline{V}_F=\overline{V}\subset W=U\cap\mathring{F}\subset F$,由 F 紧致性得到 $\overline{V}$ 的紧致性.

另一方面,$V\subset W$ 蕴含 $V=V\cap W$,故(F,τ_F)中的开集 V 也是开集 $W=U\cap\mathring{F}$ 中的开子集;但是,$W=U\cap\mathring{F}$ 是(X,τ)中的开集,故 V 也是(X,τ)中的开集. 因此得到开集 V 及其紧

致闭包 $\bar{V}$ 满足 $x\in V\subset \bar{V}_F=\bar{V}\subset U\cap \mathring{F}\subset U$；这正是正则空间的特征表示.

T_2 型局部紧致空间中的 Urysohn 引理叙述如下.

定理 3.4.20(Urysohn 引理)　设 (X,τ) 是 T_2 型局部紧致拓扑空间，$K\subset X$ 为紧致子集，$U\in\tau$ 为开集，且 $K\subset U$，则存在 X 上的实值连续函数 $f:X\to\mathbb{R}$，满足

$$f(x)\big|_{x\in K}=1,\quad f(x)\big|_{x\in \complement V}=0,$$

且 $0\leqslant f(x)\leqslant 1$，$\forall x\in X$，其中 $V\subset U$ 是 U 的一个紧致子集.

T_2 型局部紧致空间中的 Tietze 扩张定理叙述如下.

定理 3.4.21　设 (X,τ) 是 T_2 型局部紧致拓扑空间，$K\subset X$ 为紧致集，若 $f:K\to\mathbb{R}$ 是 K 上的实值连续函数，则存在 X 上的连续函数 $F:X\to\mathbb{R}$，满足 $F(x)\big|_{x\in K}=f(x)$.

3. 拓扑空间中的其他几种紧性

值得提到的是拓扑空间中的另外几种紧性.

定义 3.4.11(Lindelöff 紧性)　设 (X,τ) 是拓扑空间，若 X 的每个开覆盖，都存在可数子覆盖，则称 X 是 Lindelöff 紧的，并称 X 为 Lindelöff 空间.

定理 3.4.22(Lindelöff 定理)　设 (X,τ) 是满足第二可数公理的拓扑空间，则 X 是 Lindelöff 空间；若 (X,τ) 是 Lindelöff 度量空间，则 X 满足第二可数公理.

定义 3.4.12(可数紧性、序列紧性、列紧性)　设 (X,τ) 为拓扑空间.

(1) 若 X 的每个可数开覆盖都存在有限子覆盖，则称 (X,τ) 为**可数紧(countable compact)空间**.

(2) 若 X 的每个无限序列都存在收敛的子序列(且极限属于 X)，则称 (X,τ) 为**序列紧(sequentially compact)空间**.

(3) 若 X 的每个无限子集都存在聚点(且属于 X)，则称 (X,τ) 为**列紧空间**.

对于几种紧性的关系，我们有下列结论：

定理 3.4.23　设 (X,τ) 是拓扑空间，则

(1) 若 (X,τ) 是一般的拓扑空间，则

$$\text{紧致性}\Rightarrow\text{可数紧性}\Rightarrow\text{列紧性};$$
$$\Uparrow$$
$$\text{序列紧性}$$

(2) 若 (X,τ) 是满足第一可数公理(第二可数公理更成立)的拓扑空间，则

$$\text{紧致性}\Rightarrow\text{可数紧性}\Leftrightarrow\text{序列紧性}\Rightarrow\text{列紧性};$$

(3) 若 (X,τ) 是满足第一可数公理的、T_1 型拓扑空间，则

$$\text{紧致性}\Rightarrow\text{可数紧性}\Leftrightarrow\text{序列紧性}\Leftrightarrow\text{列紧性};$$

(4) 若 (X,ρ) 是度量空间，则

$$\text{紧致性}\Leftrightarrow\text{可数紧性}\Leftrightarrow\text{序列紧性}\Leftrightarrow\text{列紧性}.$$

3.4.4 拓扑线性空间

在前面的研究中，对于拓扑空间(X,τ)，没有提到集合X的代数结构. 现在，可考虑在拓扑空间(X,τ)中同时引进代数结构的一种空间，称为拓扑线性空间（也称为拓扑向量空间）.

定义 3.4.13（拓扑线性空间） 设在集合X中赋予加法运算$+$，数乘运算$\alpha\cdot$，$\alpha\in\mathbb{F}$，使其成为一个线性空间$(X,+,\alpha\cdot)$；又赋予以拓扑结构τ，使其成为拓扑空间(X,τ)，并且

(1) 加法运算$+$在拓扑τ之下是连续的，亦即，映射$(x,y)\to x+y$连续；

(2) 数乘运算$\alpha\cdot$在拓扑τ之下是连续的，亦即，映射$(\alpha,x)\to\alpha x$连续，

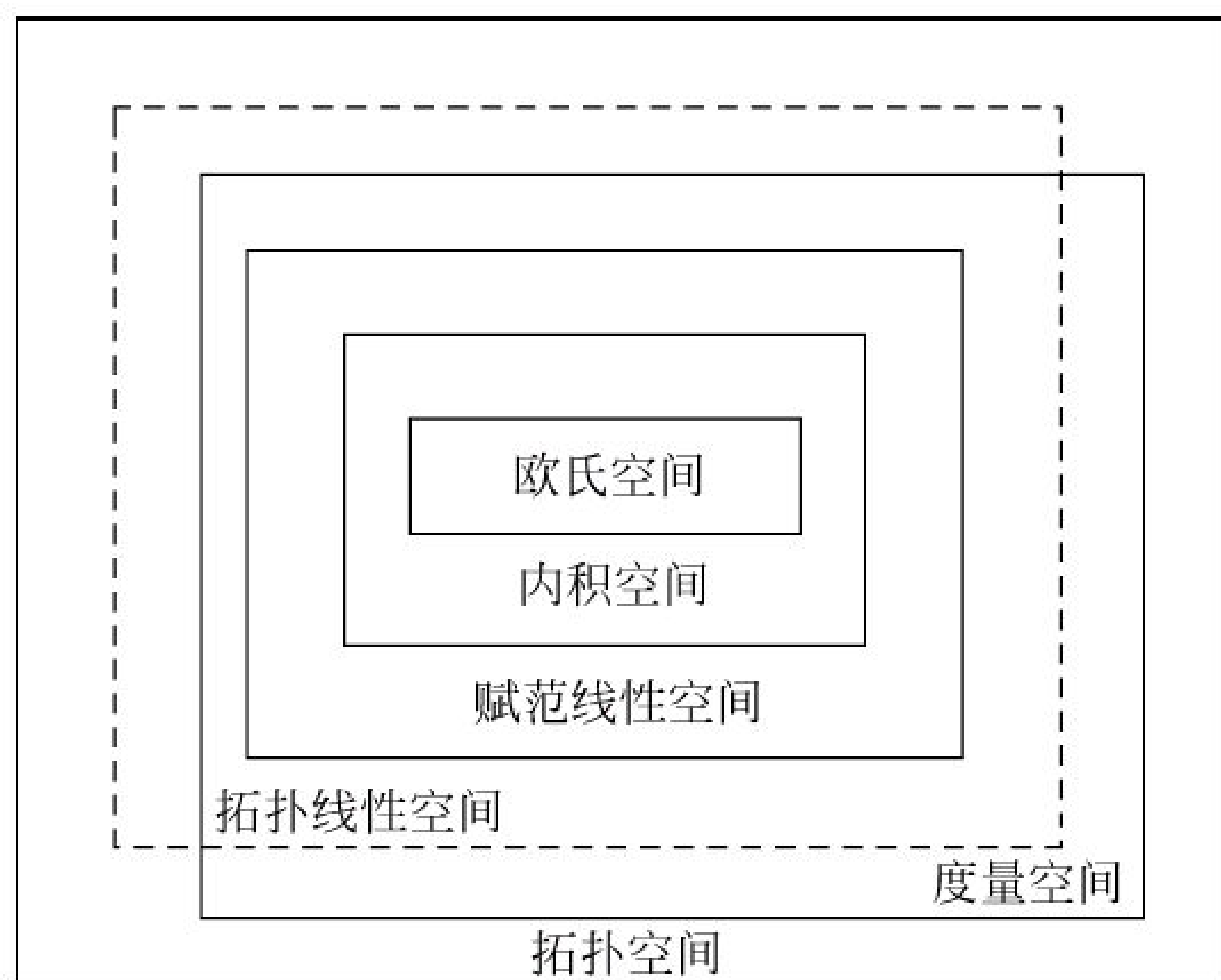
则称X为**拓扑线性空间（topological linear space）**，常记为$(X,+,\alpha\cdot,\tau)$.

在有些参考书上，还对拓扑线性空间加上“(X,τ)是Hausdorff空间”的条件.

赋范线性空间$(X,\|\cdot\|)$、内积空间$(X,(\cdot,\cdot))$、欧氏空间$\mathbb{R}^n$都是拓扑线性空间.

关于“拓扑空间”包含关系，我们有一个简单的示意图：

欧氏空间⊂内积空间⊂赋范线性空间⊂拓扑线性空间、度量空间⊂拓扑空间



数学科学和自然科学邻域中遇到的空间远远不止这些，读者会在本书的后几章中以及今后的科学研究中遇到更多的、各种各样的空间.

习题 3

1. 设(X,τ)为拓扑空间，$A\subset X$为X的子集，试证：

(1) 闭包$\overline{A}=A\cup A'$的等价定义：“X中包含集合A的所有闭集$F\subset X$的交$\bigcap\limits_{A\subset F}F$，称为$A$的**闭包**，记为$\overline{A}=\bigcap\limits_{A\subset F}F$”；

(2) 内部 $\mathring{A}$ 的等价定义："集 A 内部 $\mathring{A}$ 是包含在 A 中的最大开集 $\mathring{A}=\bigcup\limits_{\substack{G\subset A\\ G为开集}} G$"；

(3) $\overline{A}=\mathring{A}\cup\partial A$.

2. 试证导集的性质：(1)$A\subset B\Rightarrow A'\subset B'$；(2)$(A\cup B)'=A'\cup B'$；(3)$(A')'=A\cup A'$.

3. 试证闭集的性质：(1)$A\subset B\Rightarrow \overline{A}\subset\overline{B}$；(2)$\overline{A\cup B}=\overline{A}\cup\overline{B}$；(3) $\overline{\overline{A}}=\overline{A}$.

4. 试证：任何度量空间都满足第一可数性公理.

5. 任何满足第二可数性公理的拓扑空间都是可分的，试证明之.

6. 设(X,ρ)为度量空间，试证(X,ρ)具有可数的拓扑基等价于(X,ρ)是可分的.

7. 试证：$\mathbb{R}^n$ 中的"球" $B(x,r)=\left\{y\in X:\sqrt{\sum\limits_{j=1}^{n}(x_j-y_j)^2}<r\right\}$ 与"长方体"$\widetilde{B}(x,a_1,\cdots,a_n)=\{x\in\mathbb{R}^n:|x_1|<a_1,\cdots,|x_n|<a_n\}$所决定的$\mathbb{R}^n$ 上的拓扑$(\mathbb{R}^n,\tau)$与$(\mathbb{R}^n,\widetilde{\tau})$"等价"，亦即证明：$\forall O\in\tau\Rightarrow\exists\widetilde{O}\in\widetilde{\tau}$, s. t. $\widetilde{O}\subset O$；反之亦然.

8. 对于(X,τ)，(Y,υ)分别是赋范线性空间$(X,\|x\|)$与度量空间(Y,ρ)，试证定理 3.3.1.（提示：利用 A，B 之间的映射 $f:A\to B$ 成立逆像公式：$f^{-1}(A\cup B)=f^{-1}(A)\cup f^{-1}(B)$，$f^{-1}(A\cap B)=f^{-1}(A)\cap f^{-1}(B)$，$f^{-1}(A\backslash B)=f^{-1}(A)\backslash f^{-1}(B)$.）

9. 设(X,ρ)是度量空间，$\{x_j\}$是 X 中的一个序列，$x\in X$ 是 X 中的一点，试证以下条件等价：

(1) 序列$\{x_j\}$收敛到 x；

(2) 任给 $\varepsilon>0$，存在 $N\in\mathbb{Z}^+$，使当 $j>N$ 时，有 $\rho(x_j,x)<\varepsilon$；

(3) $\lim\limits_{j\to+\infty}\rho(x_j,x)=0$.

10. 试证：连续映射的复合映射是连续的.

11. 设 $X_0\subset X$ 是(X,τ)的开子空间（亦即，(X_0,τ_0)是(X,τ)的子空间，且 X_0 是 X 的开子集）. 若 $A\subset X_0$ 是 X_0 的开子集，试证：A 也是 X 的开子集.

12. 设(X,τ)，(Y,υ)为两个拓扑空间，$X_0\subset X$ 为 X 的子集、$Y_0\subset Y$ 为 Y 的子集. 两个子空间(X_0,τ_0)，(Y_0,υ_0)也构成积空间. 试证：作为 $X\times Y$ 的子集 $X_0\times Y_0$，由积空间$(X\times Y,\tau\times\upsilon)$的拓扑得到的积拓扑$(X_0\times Y_0,\tau_0\times\upsilon_0)$与由两个子拓扑空间$(X_0,\tau_0)$与$(Y_0,\upsilon_0)$所构成的积拓扑具有相同的拓扑基.

13. 试证积拓扑是使得每个投影映射为连续的最粗拓扑；商拓扑是使得商映射为连续的最精拓扑.

14. 设 $X=[0,1)\cup\{2\}$，试给出 X 在一维欧氏空间$\mathbb{R}$ 的拓扑基 $\mathfrak{B}=\{(a,b):-\infty<a<b<+\infty\}$之下的子拓扑的拓扑基.

15. 设(X,τ)，(Y,υ)为两个拓扑空间，$A\subset X$ 为 X 的闭子集、$B\subset Y$ 为 Y 的子集. 试证：$A\times B$ 是 $X\times Y$ 中的闭集.

16. 设(X,τ)是正则空间，当且仅当对于 X 的任意点 $x\in X$ 与包含 x 的任一开集 $U\in\tau$，存在一个开集 $O\in\tau$，使得 $x\in O\subset\overline{O}\subset G$. 试证明之.

17. 设(X,τ)是拓扑空间，若 X 是不连通的，试证：(1)X 中存在两个非空闭集 A，$B\subset X$，使得 $A\cap B=\varnothing$，$A\cup B=X$；(2)X 中存在两个非空开集 G，$O\subset X$，使得 $G\cap O=\varnothing$，$G\cup O=X$.

18. 试证：度量空间是 T_2 空间（Hausdorff 型），也是正规空间.

19. 试证：设(X,τ)是 T_2 型空间，若两紧致子集 A，B 互不相交，则它们是分离的.

20. 试证紧致子集的等价定义：拓扑空间(X,τ)中，子集 $A\subset X$ 为紧致的，当且仅当 A 的任意开覆盖

都存在有限子覆盖.

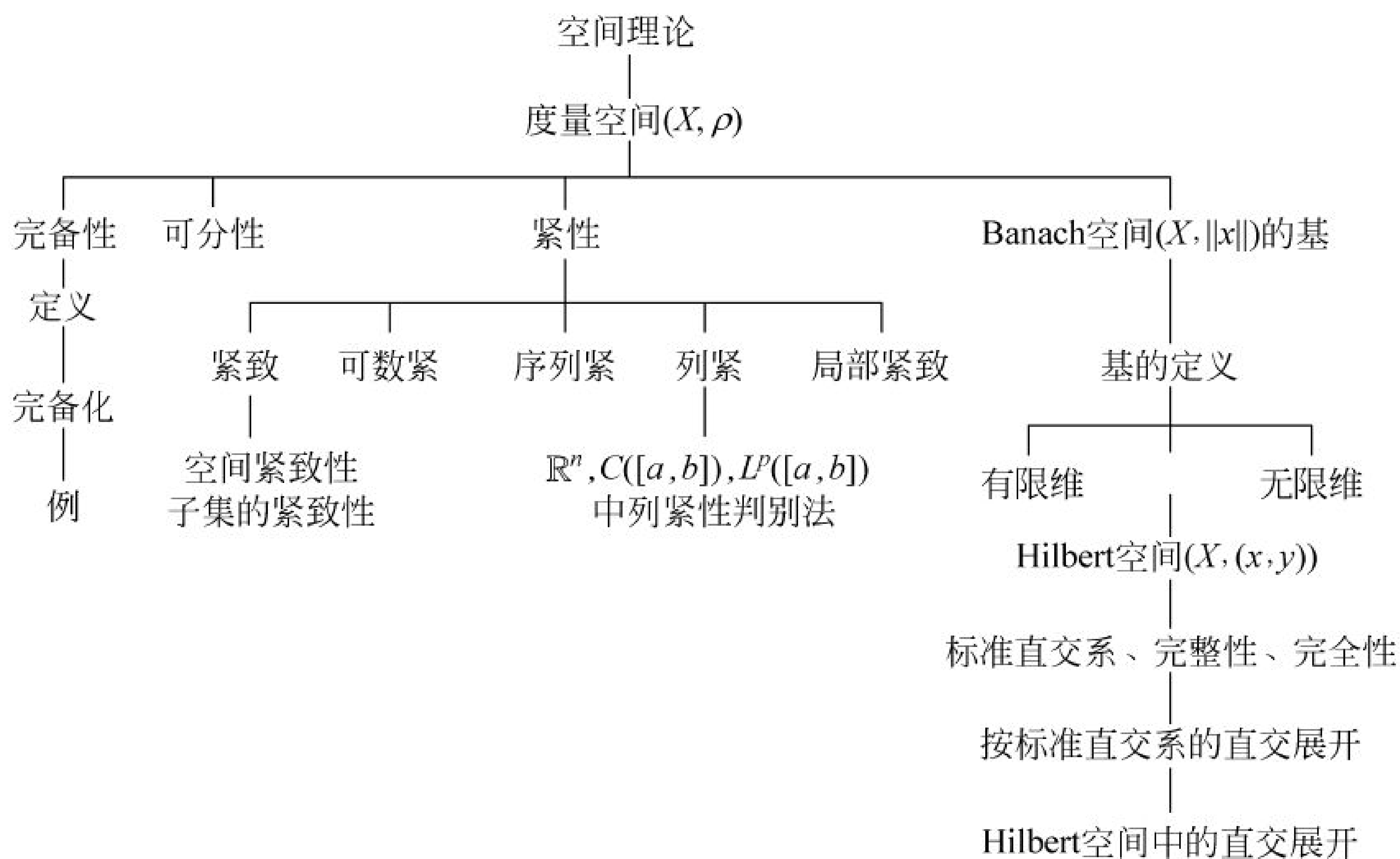
21. 试证：紧致度量空间中的**闭集套原理**：设(X,ρ)为紧致度量空间，对于任意非空递降闭集序列$\{F_n\}$：$F_1\supset F_2\supset\cdots\supset F_n\supset\cdots$，若直径$d(F_n)=\sup\{\rho(x,x')：x,x'\in F_n\}\to 0$，则存在惟一$x\in X$，使得$x\in\bigcap\limits_{n=1}^{+\infty}F_n$.（提示：利用紧致性的有限交性质——拓扑空间$(X,\rho)$中的子集族$\mathfrak{F}=\{F\subset X\}$若满足“每个有限子族都有非空交”，则称$\mathfrak{F}$具有**有限交性质**. 而紧致空间的充要条件是：具有有限交性质的闭集族都有非空的交集.）

22. 试证：若拓扑空间满足第二可数公理，则X为 Lindelöff 空间，但反之不然；进而，若(X,ρ)为度量空间，则X满足第二可数公理，当且仅当X为 Lindelöff 空间.

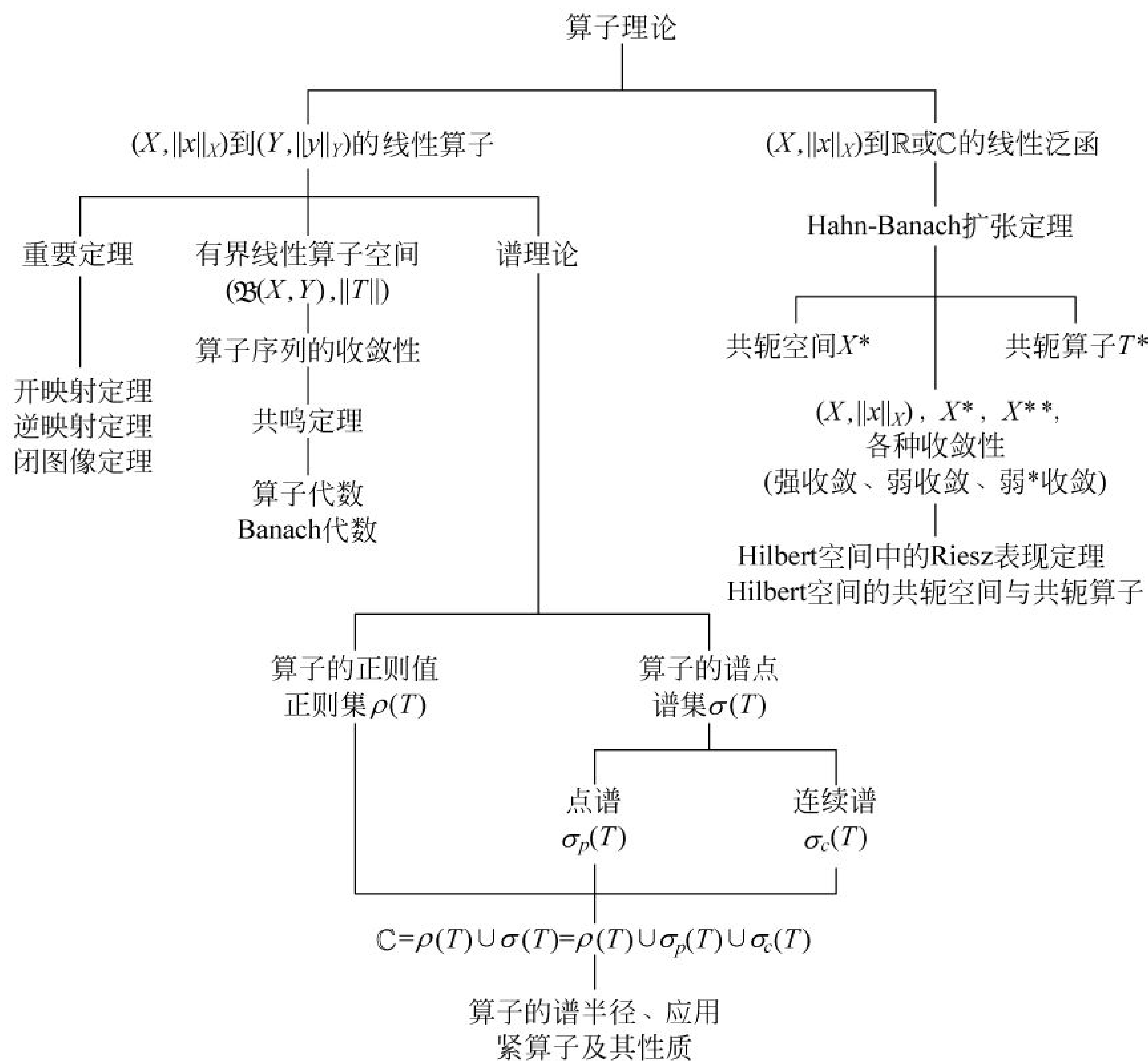
第4章 泛函分析基础

泛函分析是现代数学的一个分支，它研究既具有代数运算，又具有拓扑结构的重要集合，例如赋范线性空间、内积空间、拓扑线性空间等；同时也研究这些集合之间映射的性质，并将研究结果付诸应用. 度量空间虽然不一定有运算结构，但它的拓扑结构简单明了，有直观的几何意义，因此也是泛函分析研究的重要对象之一. 泛函分析的主要研究内容可用下图简单归纳：

关于空间理论，以度量空间为主，讨论空间的完备性、可分性与紧性，并且以 Bananch 空间、Hilbert 空间作为特例. 对于“基”，是在 Banach 空间中讨论，而“直交展开”，则对 Hilbert 空间进行讨论. 具体理论框架见下图.



算子理论是泛函分析研究对象中最重要的部分之一，以度量空间上的算子理论为基础，给出本质性的概念、基础性的定理、技巧性的证明与典型性的方法，从而得到赋范线性空间与内积空间上的线性算子与线性泛函的重要性质，作为应用学科的重要工具，是近代应用数学的必修基础.具体理论框架如下图所示.



本章主要参考文献是[6],[17].

4.1 度量空间理论

4.1.1 度量空间的完备化

1. 完备性

作为一个特殊的拓扑空间，度量空间(X,ρ)的定义已由定义3.1.1给出，并且给出几个重要的例子，如$\mathbb{R}^n$，l^2，$C([a,b])$，$L^1(E)$，$L^2(E)$等. 许多相关的重要概念，如度量空间(X,ρ)中的开集、邻域、内点、外点，闭集、聚点、孤立点、闭包，映射的连续性、序列的极限，空间的分

离性、连通性、紧性等，对于所有度量空间都成立. 为讨论度量空间更进一步的重要性质，本小节从序列的收敛性开始.

1）度量空间中的序列、收敛性

定义 4.1.1（序列的收敛性） 设(X,ρ)是度量空间，$\{x_n\}_{n\in\mathbb{Z}^+}\subset X$ 是 X 中的一个序列（由定义 3.3.3，它是$\mathbb{Z}^+$到(X,ρ)的一个映射），如果存在一个元 $a\in X$，使得当 $n\to+\infty$时，有$\rho(x_n,a)\to 0$，则称序列$\{x_n\}_{n\in\mathbb{Z}^+}$收敛于 a（或收敛到 a），记为

$$\lim_{n\to+\infty} x_n = a,$$

或简记为 $x_n\to a$，称 a 为$\{x_n\}_{n\in\mathbb{Z}^+}$的极限. 在不发生混淆的情况下，序列简记为$\{x_n\}$.

显然，当拓扑空间(X,τ)为度量空间(X,ρ)时，定义 3.3.3 与定义 4.1.1 所述是一致的. 用 ε-δ 语言描述为

$$\lim_{n\to+\infty} x_n = a \Leftrightarrow \forall\varepsilon>0,\ \exists N\in\mathbb{N},\ \text{s. t. } n>N \text{ 时},\ \rho(x_n,a)<\varepsilon.$$

关于极限，我们有如下性质：

定理 4.1.1 设(X,ρ)是度量空间，$\{x_n\}\subset X$ 是 X 中的序列.

（1）若$\{x_n\}$收敛，则其极限是惟一的；

（2）若$\{x_n\}$收敛，且极限为 $a\in X$，则其任一子列$\{x_{n_k}\}$都收敛到 a；

（3）若$\{x_n\}$收敛到 $a\in X$，则对于任意 $b\in X$，数列$\{\rho(x_n,b)\}$在$\mathbb{R}$中是有界的.

证 （1）设 $\lim\limits_{n\to+\infty}x_n=a$，$\lim\limits_{n\to+\infty}x_n=b$，且 $a\neq b$，则 $\lim\limits_{n\to+\infty}\rho(x_n,a)=\lim\limits_{n\to+\infty}\rho(x_n,b)=0$，故$\forall\varepsilon>0$，$\exists N>0$，s. t. $n>N$ 时，$\rho(x_n,a)<\varepsilon$ 且 $\rho(x_n,b)<\varepsilon$. 于是，$n>N$ 时，

$$\rho(a,b)\leqslant\rho(a,x_n)+\rho(x_n,b)<2\varepsilon,$$

由于 $\varepsilon>0$ 是任意的，故 $\rho(a,b)=0$，从而 $a=b$.（1）得证.

（2）设 $\lim\limits_{n\to+\infty}x_n=a$，故 $\forall\varepsilon>0$，$\exists N>0$，s. t. $n>N$ 时，$\rho(x_n,a)<\varepsilon$；因此对于任一子列$\{x_{n_k}\}$，当 $n_k>N$ 时，$\rho(x_{n_k},a)<\varepsilon$，此即 $\lim\limits_{k\to+\infty}x_{n_k}=a$，（2）得证.

（3）设 $\lim\limits_{n\to+\infty}x_n=a$，则$\forall\varepsilon>0$，$\exists N>0$，s. t. $n>N$ 时，$\rho(x_n,a)<\varepsilon\Rightarrow\forall b\in X$，$\rho(a,b)=\gamma$，当 $n>N$ 时，$\rho(x_n,b)\leqslant\rho(x_n,a)+\rho(a,b)<\varepsilon+\gamma\Rightarrow$ 令 M 为 $\rho(x_1,b),\cdots,\rho(x_N,b),\varepsilon+\gamma$ 中最大者，则 $\rho(x_n,b)<M(n\in\mathbb{Z}^+)$，故$\{\rho(x_n,b)\}$有界.（3）得证.

在度量空间中，连续性的充要条件可由下面定理表达.

定理 4.1.2 设(X,ρ)与(X_1,ρ_1)为两个度量空间，$f:X\to X_1$ 为 X 到 X_1 的映射，则 $f:X\to X_1$ 为连续映射，当且仅当$\forall x\in X$，序列$\{x_n\}\subset X$ 收敛到 x，蕴含序列$\{f(x_n)\}\subset X_1$ 收敛到 $f(x)$.

证 必要性 在度量空间中，映射 $f:X\to X_1$ 在 $x_0\in X$ 连续，等价地表示为

$$\forall\varepsilon>0,\ \exists\delta>0,\ \text{s. t. } \rho(x,x_0)<\delta \text{ 时},\ \rho_1(f(x),f(x_0))<\varepsilon$$

$\Rightarrow\forall\{x_n\}\subset X$，$\lim\limits_{n\to+\infty}x_n=x$，则$\forall\varepsilon>0$，$\exists N>0$，s. t. $n>N$ 时，$\rho(x_n,x)<\delta$

$\Rightarrow\rho_1(f(x_n),f(x))<\varepsilon$，也就是 $\lim\limits_{n\to+\infty}f(x_n)=f(x)$.

充分性 设 $\forall\{x_n\}\subset X$, $\lim\limits_{n\to+\infty}x_n=x$ 蕴含 $\lim\limits_{n\to+\infty}f(x_n)=f(x)$，用反证法，若 $\exists x_0\in X$，s.t. $f(x)$ 在 x_0 不连续，则 $\exists\varepsilon_0>0$，s.t. $\forall n=1,2,\cdots$，$\exists x_n\in X$，使得 $\rho(x_n,x_0)<\frac{1}{n}$，但由反证法假设 $\rho_1(f(x_n),f(x_0))\geqslant\varepsilon_0$，这与充分性假设矛盾，因此 $f:X\to X_1$ 连续.

2）度量空间中的 Cauchy 序列（基本序列）

定义 4.1.2（Cauchy 序列） 设 (X,ρ) 是度量空间，$\{x_n\}\subset X$ 是 X 中的序列，称 $\{x_n\}$ 为 X 中 **Cauchy 序列（基本序列）**，若对于任意 $\varepsilon>0$，存在 $N>0$，使得 $m,n>N$ 时，$\rho(x_m,x_n)<\varepsilon$.

3）度量空间的完备性与完备化

定义 4.1.3（完备度量空间） 设 (X,ρ) 是度量空间，如果 (X,ρ) 中的任一 Cauchy 序列都收敛到 X 中的一个元 $x\in X$，则称 (X,ρ) 是**完备度量空间（complete metric space）**.

度量空间 (X,ρ) 的子集 $A\subset X$，如果作为子空间是完备的，则称 A 为 (X,ρ) 的**完备子集（complete subset）**.

例 4.1.1 设 $\mathbb{Q}$ 为有理数集，在距离 $\rho(x,y)=|x-y|$ 之下成为一个度量空间，且是 $(\mathbb{R},\rho)$ 的子空间. 但在距离 $\rho(x,y)=|x-y|$ 之下，$\mathbb{Q}$ 是不完备的.

事实上，任一有理数序列 $\{r_n\}$，$r_n\in\mathbb{Q}$，若在距离 $\rho(x,y)=|x-y|$ 下收敛，则其极限可以是有理数 $r\in\mathbb{Q}$，也可以是无理数 $s\notin\mathbb{Q}$. 例如，对于有理数序列 $r_n=1+\frac{1}{n}$，有

$$\lim_{n\to+\infty}r_n=\lim_{n\to+\infty}\left(1+\frac{1}{n}\right)=1\in\mathbb{Q};$$

但是，对于有理数序列 $(r_n)^n=\left(1+\frac{1}{n}\right)^n$，却有

$$\lim_{n\to+\infty}(r_n)^n=\lim_{n\to+\infty}\left(1+\frac{1}{n}\right)^n=\mathrm{e}\notin\mathbb{Q}.$$

例 4.1.2 设 $\mathbb{R}$ 为实数集，在距离 $\rho(x,y)=|x-y|$ 之下成为一个度量空间 $(\mathbb{R},\rho)$，并且是完备度量空间.

事实上，对于 $(\mathbb{R}^1,|x-y|)$ 中的任一 Cauchy 序列 $\{x_n\}\subset\mathbb{R}^1$，取 $\varepsilon=1$，$\exists N\in\mathbb{N}$，当 $n,m>N$ 时，$|x_n-x_m|<1$；于是，取定 $m=m_0>N$，则 $n>N$ 时，$|x_n-x_{m_0}|<1$，故

$$|x_n|<|x_n-x_{m_0}|+|x_{m_0}|<1+|x_{m_0}|,\quad\forall n>N,$$

由此，$|x_n|\leqslant\max\{|x_1|,\cdots,|x_{m_0}|,|x_{m_0}|+1\}\equiv M$，$\forall n\in N$ 成立，所以，$\{x_n\}$ 是 $\mathbb{R}^1$ 中的有界序列. 由实数集中闭区间的 Bolzano-Weierstrass 定理得到：对于 $\mathbb{R}^1$ 中的有界序列 $\{x_n\}$，必存在收敛的子序列 $\{x_{n_k}\}$，使得 $\lim\limits_{k\to+\infty}x_{n_k}=x\in[-M,M]$. 下面证明 $\{x_n\}$ 也收敛到同一个 x.

事实上，$\forall \varepsilon>0$，$\exists N_1\in\mathbb{N}$，s. t. $k>N_1$ 时，$|x_{n_k}-x|<\frac{\varepsilon}{2}$；此时必有 $n_k\geqslant k$（子列的定义）．另一方面，因 $\{x_n\}$ 是 Cauchy 序列，故 $\exists N_2\in\mathbb{N}$，s. t. $n,m>N_2$ 时，$|x_n-x_m|<\frac{\varepsilon}{2}$；于是，令 $N=\max\{N_1,N_2\}$，因此 $k>N$；从而 $n_k>N$ 时，$|x_{n_k}-x|<\frac{\varepsilon}{2}$；于是，$n>N$ 时，有

$$|x_n-x|\leqslant|x_n-x_{n_k}|+|x_{n_k}-x|<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon,$$

故得到 $\lim\limits_{n\to+\infty}x_n=x\in\mathbb{R}^1$．$\mathbb{R}^1$ 的完备性得证．

4）常用的完备度量空间

（1）$(\mathbb{R}^n,\rho(x,y))$．

$$\mathbb{R}^n=\{x=(x_1,x_2,\cdots,x_n):x_j\in\mathbb{R},\ j=1,2,\cdots,n\},\quad \rho(x,y)=\sqrt{\sum_{j=1}^{n}(x_j-y_j)^2}.$$

（2）$(l^p,\rho(x,y))$，$1\leqslant p<+\infty$．

$$l^p=\left\{x=(x_1,x_2,\cdots):\sum_{j=1}^{+\infty}|x_j|^p<+\infty\right\},\quad \rho(x,y)=\left(\sum_{j=1}^{\infty}|x_j-y_j|^p\right)^{\frac{1}{p}}.$$

（3）$(l^\infty,\rho(x,y))$．

$$l^\infty=\{x=(x_1,x_2,\cdots):|x_j|<+\infty,j\in\mathbb{Z}^+\},\quad \rho(x,y)=\sup_{j\in\mathbb{Z}^+}|x_j-y_j|.$$

（4）$(s,\rho(x,y))$．

$$s=\{x=(x_1,x_2,\cdots):x_j\in\mathbb{R}\},\quad \rho(x,y)=\sum_{j=1}^{\infty}\frac{1}{2^j}\frac{|x_j-y_j|}{1+|x_j-y_j|}.$$

（5）$(S([a,b]),\rho(x,y))$．

$$S([a,b])=\{f:[a,b]\text{ 上 a. e. 有限的可测函数}\},\rho(x,y)=\int_{[a,b]}\frac{|f(t)-g(t)|}{1+|f(t)-g(t)|}\mathrm{d}t.$$

（6）$(C([a,b]),\rho(x,y))$．

$$C([a,b])=\{x(t):[a,b]\text{ 上的连续函数}\},\ \rho(x,y)=\max_{t\in[a,b]}|x(t)-y(t)|.$$

（7）$(C^k([a,b]),\rho(x,y))$，$k\in\mathbb{N}$．

$C^k([a,b])=\{x(t):x^{(k)}(t)\in C([a,b])\}$，约定 $C^{(0)}([a,b])=C([a,b])$，

$$\rho(x,y)=\sum_{j=0}^{k}\max_{t\in[a,b]}|x^{(j)}(t)-y^{(j)}(t)|.$$

（8）$(C^\infty([a,b]),\rho(x,y))$．

$C^\infty([a,b])=\{x(t):x^{(k)}(t)\in C([a,b]),k\in\mathbb{N}\}$，

$$\rho(x,y)=\sum_{n=0}^{+\infty}\frac{1}{2^n}\frac{\max\limits_{t\in[a,b]}|x^{(n)}(t)-y^{(n)}(t)|}{1+\max\limits_{t\in[a,b]}|x^{(n)}(t)-y^{(n)}(t)|}.$$

(9) $(L^p(E),\rho(x,y))$，$1\leqslant p<+\infty$，$E\subset\mathbb{R}$ 为 Lebesgue 可测集.

$$L^p(E)=\left\{x(t):\int_E|x(t)|^p\mathrm{d}t<+\infty\right\},\quad \rho(x,y)=\left(\int_E|x(t)-y(t)|^p\mathrm{d}t\right)^{\frac{1}{p}}.$$

(10) $(L^\infty(E),\rho(x,y))$，$E\subset\mathbb{R}$ 为 Lebesgue 可测集.

$$L^\infty(E)=\{x(t):\operatorname*{esssup}_{t\in E}|x(t)|\overset{\text{def}}{=}\inf_{\substack{m(E_0)=0\\E_0\subset E}}\sup_{t\in E\setminus E_0}|x(t)|<+\infty\},$$

$$\rho(x,y)=\operatorname*{esssup}_{t\in E}|x(t)-y(t)|.$$

为讨论度量空间的完备化，对照一般拓扑空间(X,τ)中稠密性的定义，度量空间中则有如下定义.

定义 4.1.4(稠密子集) 设 $A\subset X$，$B\subset X$ 是度量空间(X,ρ)的子集，称 A 在 B 中**稠密**，若 $B\subseteq\overline{A}$.

注 1 对于整个空间 X 而言，若 $X=\overline{A}$，则称 A 是 X 的稠密子集.

注 2 度量空间(X,ρ)中稠密性的意义：$B\subseteq\overline{A}$ 表示 B 中的点都是 A 的聚点，亦即

$$\forall x\in B,\ \forall\varepsilon>0\Rightarrow\exists y\in A,\ \text{s. t.}\ \rho(x,y)<\varepsilon;$$

或等价地，有

$$\forall x\in B\Rightarrow\exists\{y_n\}\subset A,\ \text{s. t.}\ \lim_{n\to+\infty}\rho(x,y_n)=0.$$

定义 4.1.5(度量空间的完备化) 对于度量空间(X,ρ)，包含(X,ρ)的最小完备度量空间(X_0,ρ_0)称为(X,ρ)的**完备化度量空间**，简称完备度量空间(X_0,ρ_0)为度量空间(X,ρ)的**完备化(completion)**.

此定义等价于：对于度量空间(X,ρ)，若存在完备度量空间(X_0,ρ_0)，使得(X,ρ)与(X_0,ρ_0)的一个稠密子集$\widetilde{X}_0$ **等距同构(equidistant isomorphism)**；亦即，若存在完备度量空间(X_0,ρ_0)及其稠密子集$\widetilde{X}_0\subset X_0$，$X_0=\overline{\widetilde{X}_0}$，并且存在映射 $T:X\to\widetilde{X}_0$，s. t.

$$\rho_0(Tx,Ty)=\rho(x,y),\quad\forall x,y\in X,$$

其中映射 $T:X\to\widetilde{X}_0$ 称为 X 到$\widetilde{X}_0$ 的等距同构映射，即 T 是 X 到$\widetilde{X}_0$ 的双方单值、双方连续，且保持距离 $\rho_0(Tx,Ty)=\rho(x,y)$的映射，则(X_0,ρ_0)称为(X,ρ)的**完备化**，映射 T 也简称为**等距同构**.

定理 4.1.3 设(X,ρ)是度量空间，则(X,ρ)必存在其完备化(X_0,ρ_0)；进而，在等距意义下，度量空间(X,ρ)的完备化是惟一的.

证 根据定义，需要构造一个完备度量空间(X_0,ρ_0)，并构造(X_0,ρ_0)的稠密子集$\widetilde{X}_0\subset X_0$，$X=\overline{\widetilde{X}_0}$，使得$(\widetilde{X}_0,\rho_0)$与$(X,\rho)$等距同构. 具体方法是利用等价类来构造$(X,\rho)$的完备

化.证明分为以下六步.

(1) 定义等价关系～：对于(X,ρ)中的 Cauchy 序列$\{x_n\},\{y_n\}\subset X$,若$\lim\limits_{n\to+\infty}\rho(x_n,y_n)=0$,则称此二 Cauchy 序列彼此等价,记为$\{x_n\}\sim\{y_n\}$. 显然这是等价关系；将每个等价类记为$\xi=\{x_n\},\eta=\{y_n\},\cdots$,等价类的全体记为

$$X_0=\{\xi,\eta,\zeta,\cdots\}=\{\{x_n\},\{y_n\},\{z_n\},\cdots\}.$$

(2) 在$X_0=\{\xi,\eta,\zeta,\cdots\}$中定义距离,使其成为度量空间：对于两个元$\xi=\{x_n\}\in X_0$与$\eta=\{y_n\}\in X_0$,定义其间的“距离”为

$$\rho_0(\xi,\eta)=\lim_{n\to+\infty}\rho(x_n,y_n). \tag{4.1.1}$$

我们证明,此ρ_0为$X_0=\{\xi,\eta,\zeta,\cdots\}$上的一个距离. 亦即证明以下三点：

① (4.1.1)式定义的极限存在,且为非负实数. 这只要证明实数序列$\{\alpha_n\}=\{\rho(x_n,y_n)\}$是一个 Cauchy 序列,再根据实数的连续性,存在极限$\lim\limits_{n\to+\infty}\alpha_n=\lim\limits_{n\to+\infty}\rho(x_n,y_n)\in\mathbb{R}$,且由于$\alpha_n\geqslant 0$得到$\lim\limits_{n\to+\infty}\alpha_n=\alpha\geqslant 0$.

事实上,实数序列$\{\alpha_n\}=\{\rho(x_n,y_n)\}$是 Cauchy 序列,因为由距离的三点不等式得

$$\begin{aligned}|\alpha_n-\alpha_m|&=|\rho(x_n,y_n)-\rho(x_m,y_m)|\\&\leqslant|\rho(x_n,y_n)-\rho(x_m,y_n)|+|\rho(x_m,y_n)-\rho(x_m,y_m)|\\&\leqslant\rho(x_n,x_m)+\rho(y_n,y_m);\end{aligned}$$

于是,由$\xi=\{x_n\},\eta=\{y_n\}$为 Cauchy 序列,故$\forall\,\varepsilon>0$,$\exists\,N>0$, s. t. $n,m>N$时,有

$$\rho(x_n,x_m)<\frac{\varepsilon}{2},\quad\rho(y_n,y_m)<\frac{\varepsilon}{2},$$

从而得到$\forall\,\varepsilon>0$,$\exists\,N>0$, s. t. $n,m>N$时,$|\alpha_n-\alpha_m|<\varepsilon$. 此即$\{\alpha_n\}=\{\rho(x_n,y_n)\}$是$\mathbb{R}$中的 Cauchy 序列.

② (4.1.1)式中的极限不依赖于等价类ξ中 Cauchy 序列$\{x_n\}$的选择. 亦即要证明：若$\{x_n\},\{x'_n\}\in\xi,\{y_n\},\{y'_n\}\in\eta$,则

$$\rho_0(\xi,\eta)=\lim_{n\to+\infty}\rho(x_n,y_n)=\lim_{n\to+\infty}\rho(x'_n,y'_n).$$

事实上,由

$$\begin{aligned}|\rho(x_n,y_n)-\rho(x'_n,y'_n)|&\leqslant|\rho(x_n,y_n)-\rho(x'_n,y_n)|+|\rho(x'_n,y_n)-\rho(x'_n,y'_n)|\\&\leqslant\rho(x_n,x'_n)+\rho(y_n,y'_n),\end{aligned}$$

但因$\{x_n\},\{x'_n\}\in\xi$蕴含$\rho(x_n,x'_n)\to 0$与$\{y_n\},\{y'_n\}\in\eta$蕴含$\rho(y_n,y'_n)\to 0$,故上式右边趋向于零,从而得到$\lim\limits_{n\to+\infty}\rho(x_n,y_n)=\lim\limits_{n\to+\infty}\rho(x'_n,y'_n)$.

③ (4.1.1)式中定义的$\rho_0(\xi,\eta)=\lim\limits_{n\to+\infty}\rho(x_n,y_n)$满足距离的三个条件(证明留作习题).

于是,$X_0=\{\xi,\eta,\zeta,\cdots\}$在距离$\rho_0(\xi,\eta)=\lim\limits_{n\to+\infty}\rho(x_n,y_n)$之下成为一个度量空间$(X_0,\rho_0)$.

(3) **在$X_0=\{\xi,\eta,\zeta,\cdots\}$中定义一个稠密子集$\widetilde{X}_0\subset X_0$**,亦即$X_0=\overline{\widetilde{X}_0}$：事实上,对$x\in X$,显然$x_n=x,n=1,2,\cdots$是 Cauchy 序列,称为**常驻序列**,记为$\tilde{\xi}_x\equiv\{x,x,\cdots\}\in X_0$. **记常驻序**

列的全体所成的集为$\widetilde{X}_0=\{\widetilde{\xi}_x: x\in X\}\subset X_0$. 下面证 $\widetilde{X}_0$ 在 X_0 中稠密,即证 $X_0=\overline{\widetilde{X}_0}$.

推理如下. $\forall \xi=\{x_n\}\in X_0 \Rightarrow \forall \varepsilon>0, \exists N>0$, s. t. $n,m>N$ 时,$\rho(x_n,x_m)<\varepsilon \Rightarrow \forall j>N, \exists \widetilde{\xi}_j\equiv\widetilde{\xi}_{x_j}=\{x_j,\cdots,x_j,\cdots\}\in\widetilde{X}_0$, s. t. $\rho_0(\widetilde{\xi}_j,\xi)=\lim\limits_{n\to+\infty}\rho(x_j,x_n)\leqslant\varepsilon \Rightarrow \lim\limits_{j\to+\infty}\rho_0(\widetilde{\xi}_j,\xi)=0$ $\Rightarrow \forall \xi=\{x_n\}\in X_0, \exists \widetilde{\xi}_j\in\widetilde{X}_0$, s. t. $\lim\limits_{j\to+\infty}\rho_0(\widetilde{\xi}_j,\xi)=0 \Rightarrow X_0=\overline{\widetilde{X}_0}$.

(4) (X_0,ρ_0)是完备度量空间:取$\{\xi_n\}\subset X_0$ 为 Cauchy 序列,欲证 $\exists \eta\in X_0$, s. t. $\lim\limits_{n\to+\infty}\rho_0(\xi_n,\eta)=0$. 事实上,$\forall \varepsilon=\frac{1}{n}, \forall \xi_n\in X_0=\overline{\widetilde{X}_0} \Rightarrow \exists \widetilde{\eta}_n\in\widetilde{X}_0$, s. t. $\rho_0(\xi_n,\widetilde{\eta}_n)<\frac{1}{n}, n=1,2,\cdots \Rightarrow \forall n$,常驻序列 $\widetilde{\eta}_n=\{y_n,\cdots,y_n,\cdots\}\in\widetilde{X}_0\subset X_0$ 是 X_0 中的一等价类 $\Rightarrow$ 等价类序列 $\{\widetilde{\eta}_n\}\subset X_0$ 满足 $\rho_0(\widetilde{\eta}_n,\widetilde{\eta}_m)=\rho(y_n,y_m)$,其中 $y_n\in X$,且 $\rho(y_n,y_m)=\rho_0(\widetilde{\eta}_n,\widetilde{\eta}_m)\leqslant\rho_0(\widetilde{\eta}_n,\xi_n)+\rho_0(\xi_n,\xi_m)+\rho_0(\xi_m,\widetilde{\eta}_m)<\frac{1}{n}+\rho_0(\xi_n,\xi_m)+\frac{1}{m} \Rightarrow$ 由 $\{\xi_n\}\subset X_0$ 为 Cauchy 序列,故 n, $m\to+\infty$时,$\rho_0(\xi_n,\xi_m)\to 0$. 于是,当 $n,m\to+\infty$时,$\lim\limits_{n,m\to\infty}\rho(y_n,y_m)=0 \Rightarrow \{y_m\}\subset X$ 是 Cauchy 序列 $\Rightarrow \{y_m\}$必属于某个等价类,从而 $\exists \eta\in X_0$, s. t. $\{y_m\}\in\eta$,且 $\rho_0(\xi_n,\eta)\leqslant\rho_0(\xi_n,\widetilde{\eta}_n)+\rho_0(\widetilde{\eta}_n,\eta)<\frac{1}{n}+\rho_0(\widetilde{\eta}_n,\eta) \Rightarrow \eta\in X_0$ 是$\{y_m\}$所属的等价类,且$\{y_n,\cdots,y_n,\cdots\}=\widetilde{\eta}_n\in\widetilde{X}_0 \Rightarrow$ $\forall n=1,2,\cdots$,有 $\rho_0(\widetilde{\eta}_n,\eta)=\lim\limits_{m\to+\infty}\rho(y_n,y_m)=0 \Rightarrow$ 当 n 充分大时,有 $\rho_0(\widetilde{\eta}_n,\eta)<\frac{1}{n} \Rightarrow$ 存在 $\eta\in X_0$, s. t. $\lim\limits_{n\to+\infty}\rho_0(\xi_n,\eta)=0 \Rightarrow (X_0,\rho_0)$的完备性得证.

(5) $\widetilde{X}_0\subset X_0$ 与 X 等距:定义映射 $T: x\to\xi_x$. 我们指出

① T 是 X 到$\widetilde{X}_0$ 的一一满射:因为$\forall x\in X, \exists!\ \xi_x\in\widetilde{X}_0$;反之,$\forall \xi\in\widetilde{X}_0$ 是常驻序列,故 $\xi=\{x_1,x_2,\cdots,x_n,\cdots\}$的每项必相等,$x_1=x_2=\cdots=x_n=\cdots\equiv x\in X$;

② T 是 X 到$\widetilde{X}_0$ 的等距映射:因为 $\rho_0(\xi,\eta)=\lim\limits_{n\to+\infty}\rho(x_n,y_n)=\lim\limits_{n\to+\infty}\rho_0(\widetilde{\xi}_x,\widetilde{\eta}_y)=\rho(x,y)$.

③ T 与 T^{-1}的连续性由等距立即得到.

(6) 在等距意义下,(X,ρ)的完备化(X_0,ρ_0)惟一确定:我们省略其证明,读者可参看[6].

于是,定理得证.

例 4.1.3 设 $X=\{x=(\xi_1,\xi_2,\cdots,\xi_k,0,0,\cdots): \xi_j\in\mathbb{R}, 1\leqslant j\leqslant k, k\in\mathbb{Z}^+\}$,它是 $l^p=\left\{x=(x_1,x_2,\cdots): \sum\limits_{j=1}^{+\infty}|x_j|^p<+\infty\right\}$ 的子空间(为什么?),但是在 l^p 的距离

$$\rho(x,y)=\left(\sum_{j=1}^{+\infty}|x_j-y_j|^p\right)^{\frac{1}{p}}$$

之下,它却不完备. 事实上,若取 X 中的序列

$$x_1=(1,0,0\cdots),\quad x_2=\left(1,\frac{1}{2},0,0\cdots\right)\quad,\cdots,\quad x_n=\left(1,\frac{1}{2},\cdots,\frac{1}{2^{n-1}},0,0,\cdots\right),\cdots,$$

易见,$\{x_n\}\subset X$ 是距离 $\rho(x,y)=\left(\sum_{j=1}^{+\infty}|x_j-y_j|^p\right)^{\frac{1}{p}}$ 之下的 Cauchy 序列,但是在子空间 X 中却无极限. 显然,X 在 l^p 中稠密,而且 l^p 是 X 的完备化.

例 4.1.4 设 $P([a,b])$为$[a,b]$上所有多项式的全体,它是空间$(C([a,b]),\rho(x,y))$的子空间,但是,在 $C([a,b])$的距离之下,$P([a,b])$却不完备,这是因为 $P([a,b])$中的序列

$$p_1(t)=1,\quad p_2(t)=1+\frac{t}{2},\quad \cdots,\quad p_n(t)=1+\frac{t}{2}+\cdots+\frac{t^{n-1}}{2^{n-1}},\quad \cdots$$

是 $C([a,b])$距离下的 Cauchy 序列,但在 $P([a,b])$中却无极限. 然而,$P([a,b])\subset C([a,b])$,且 $P([a,b])$在 $C([a,b])$中稠密,因此 $C([a,b])$是 $P([a,b])$的完备化.

例 4.1.5 $C([a,b])$在距离 $\rho(f,g)=\max\limits_{x\in[a,b]}|f(x)-g(x)|$之下是完备度量空间;$L^p([a,b])$在距离 $\rho(f,g)=\left(\int_a^b|f(x)-g(x)|^p\mathrm{d}x\right)^{\frac{1}{p}}$之下是完备度量空间. 下面给出这两个重要的度量空间完备性的证明.

(1) 对于 $C([a,b])$中的任一 Cauchy 序列$\{x_n(t):t\in[a,b]\}$,由定义得

$$\forall\varepsilon>0,\ \exists N\in\mathbb{N},\ \text{s. t. } n,m>N \text{ 时},\ \max_{t\in[a,b]}\{|x_n(t)-x_m(t)|\}<\varepsilon,$$

亦即,不等式$|x_n(t)-x_m(t)|<\varepsilon$ 对于 $t\in[a,b]$一致成立,也就是 $\lim\limits_{n\to+\infty}x_n(t)$在 $t\in[a,b]$上一致收敛,由高等数学中的结果,$\exists\, x(t)\in C([a,b])$, s. t. $\lim\limits_{n\to+\infty}x_n(t)=x(t)$,故 $C([a,b])$是完备度量空间.

(2) 对于 $L^p(E)$中的任一 Cauchy 序列$\{x_n(t):t\in E\}\subset L^p(E)$,对于 $k\in\mathbb{Z}^+$,可选取一个自然数 $n_k\in\mathbb{Z}^+$,使得 $\left\{\int_E|x_{n_k}(t)-x_{n_{k+1}}(t)|^p\mathrm{d}t\right\}^{\frac{1}{p}}<\frac{1}{2^k}$,$k=1,2,\cdots$. 对于每个具有有限测度的子集 $e\subset E$,$m(e)<+\infty$,利用 Hölder 不等式,有

$$\int_e|x_{n_k}(t)-x_{n_{k+1}}(t)|\mathrm{d}t\leqslant\left\{\int_e|x_{n_k}(t)-x_{n_{k+1}}(t)|^p\mathrm{d}t\right\}^{\frac{1}{p}}\left\{\int_e 1^q\mathrm{d}t\right\}^{\frac{1}{q}}<\frac{1}{2^k}(m(e))^{\frac{1}{q}},$$

其中$\frac{1}{p}+\frac{1}{q}=1$,因此级数 $\sum_{k=1}^{+\infty}\int_e|x_{n_k}(t)-x_{n_{k+1}}(t)|\mathrm{d}t$ 收敛. 由被积函数$|x_{n_k}(t)-x_{n_{k+1}}(t)|\geqslant 0$ 非负,根据实变函数中的正项级数的可积性定理([6]),得到

$$\sum_{k=1}^{+\infty}\int_e|x_{n_k}(t)-x_{n_{k+1}}(t)|\mathrm{d}t=\int_e\sum_{k=1}^{+\infty}|x_{n_k}(t)-x_{n_{k+1}}(t)|\mathrm{d}t.$$

于是,级数

$$|x_{n_1}(t)|+|x_{n_2}(t)-x_{n_1}(t)|+|x_{n_3}(t)-x_{n_2}(t)|+\cdots$$

在 e 中几乎处处收敛;又由 $e\subset E$ 为任一有限测度的子集,故其和函数 $x(t)$是 L 可测函数;由此,级数

$$x_{n_1}(t)+x_{n_2}(t)-x_{n_1}(t)+x_{n_3}(t)-x_{n_2}(t)+\cdots$$

在 E 上几乎处处收敛到可测函数 $x(t)$，亦即序列 $\{x_{n_k}(t)\}$ 在 E 中几乎处处收敛到 $x(t)$.

下面证明 $x(t)\in L^p(E)$，且 $\lim\limits_{n\to+\infty}\left\{\int_E |x_n(t)-x(t)|^p \mathrm{d}t\right\}^{\frac{1}{p}}=0$.

由于 $\{x_n(t):t\in E\}$ 为 Cauchy 序列，$\forall\varepsilon>0$，$\exists N\in\mathbb{N}$，s. t. $n,n_k>N$ 时，有

$$\left\{\int_E |x_{n_k}(t)-x_n(t)|^p \mathrm{d}t\right\}^{\frac{1}{p}}<\varepsilon,$$

固定 n，令 $k\to+\infty$，根据实变函数中的积分号下取极限的定理，得到

$$\begin{aligned}\lim_{k\to+\infty}\left\{\int_E |x_{n_k}(t)-x_n(t)|^p \mathrm{d}t\right\}^{\frac{1}{p}} &= \left\{\int_E \lim_{n_k\to+\infty}|x(t)-x_n(t)|^p \mathrm{d}t\right\}^{\frac{1}{p}}\\ &= \left\{\int_E |x(t)-x_n(t)|^p \mathrm{d}t\right\}^{\frac{1}{p}}<\varepsilon.\end{aligned}$$

于是，$\forall\varepsilon>0$，$\exists N\in\mathbb{Z}^+$，s. t. $n>N$ 时，$\left\{\int_E |x(t)-x_n(t)|^p \mathrm{d}t\right\}^{\frac{1}{p}}<\varepsilon$.

进而，由 $x(t)-x_n(t)\in L^p(E)$，知 $x(t)=\{x(t)-x_n(t)\}+x_n(t)\in L^p(E)$，因此得到基本序列 $\{x_n(t):t\in E\}\subset L^p(E)$ 收敛到 $x(t)\in L^p(E)$. $L^p(E)$ 完备性得证.

然而，我们指出，$C([a,b])$ 在距离

$$\rho(x,y)=\left(\int_{[a,b]}|x(t)-y(t)|^2 \mathrm{d}t\right)^{\frac{1}{2}}$$

之下，虽然是 $L^2([a,b])$ 的子空间，但却不是完备子空间.

事实上，取一点 $c\in(a,b)$，令 $x_n(t)=\arctan n(t-c)$，$t\in[a,b]$，$n\in\mathbb{Z}^+$，可以证明，

$$x_n(t)\in C([a,b])\subset L^2([a,b])；\text{且存在 } x(t)=\begin{cases}\dfrac{\pi}{2}, & c<t\leqslant b,\\ 0, & t=c,\\ -\dfrac{\pi}{2}, & a\leqslant t<b\end{cases}\in L^2([a,b])\text{ 是 } x_n(t)$$

在 $L^2([a,b])$ 距离下的极限，$\lim\limits_{n\to+\infty}\rho(x_n,x)=\lim\limits_{n\to+\infty}\left(\int_{[a,b]}|x_n(x)-x(x)|^2\mathrm{d}x\right)^{\frac{1}{2}}=0$. 因此，$\{x_n(t)\}\subset C([a,b])$，且是距离 $\rho(x,y)=\left(\int_{[a,b]}|x(t)-y(t)|^2\mathrm{d}t\right)^{\frac{1}{2}}$ 之下的 Cauchy 序列；但是，这个 $x(t)$ 却不是连续函数，所以 $C([a,b])$ 按距离 $\rho(x,y)=\left(\int_{[a,b]}|x(t)-y(t)|^2\mathrm{d}t\right)^{\frac{1}{2}}$ 是不完备子空间.

由以上例子可知，度量空间是否完备，与距离有密切关系，例如，在一个距离 ρ_1 之下，(X,ρ_1) 可以是完备的，但在另一个距离 ρ_2 之下，(X,ρ_2) 却不是完备的.

由于赋范线性空间、内积空间、欧氏空间等都是度量空间，因此本节关于完备性的讨论对它们都适合．今后，就可视度量空间、赋范线性空间、内积空间都已经完备化而成为完备空间．完备的赋范线性空间称为 **Banach 空间**，完备的内积空间称为 **Hilbert 空间**．

2. 完备度量空间的重要性质及应用

定理 4.1.4 设(X,ρ)是完备度量空间，则 X 中的任意非空闭子空间是完备子空间；反之，X 的完备子空间是 X 的闭子空间．

证 必要性 因为完备度量空间(X,ρ)中的非空闭子集 F 的任一基本序列$\{x_n\}\subset F$ 收敛于 $x_0\in X$，且 $\lim\limits_{n\to+\infty} x_n=x_0\in\overline{F}=F$，故 $x_0\in F$，所以(F,ρ)是完备子空间．

充分性 若度量空间(X,ρ)的子空间(E,ρ)是完备子空间，$E\subset X$．为证 E 是 X 的闭子集，任取 $a\in\overline{E}$，则 $\exists\{x_n\}\subset E$，s. t. $\lim\limits_{n\to+\infty} x_n=a \Rightarrow \{x_n\}\subset E$ 收敛 $\Rightarrow \{x_n\}$为基本序列 $\Rightarrow$ 因 E 完备，故 $\exists b\in E$，s. t. $\lim\limits_{n\to+\infty} x_n=b \Rightarrow$ 因度量空间为 T_2 型，故两个极限 $b\in E$ 与 $a\in\overline{E}$ 相等 $\Rightarrow a\in E \Rightarrow \overline{E}=E$．充分性得证．

完备度量空间中的**不动点（fixed point）定理**非常重要，并且有广泛的应用．

定理 4.1.5 对于完备度量空间(X,ρ)到自身的映射 $T:X\to X$，若 $\forall x,y\in X$，成立不等式

$$\rho(T(x),T(y))\leqslant\theta\rho(x,y),\tag{4.1.2}$$

其中 $0\leqslant\theta<1$ 为常数，则映射 T 在 X 中存在惟一的不动点 $x_0\in X$．亦即，存在惟一的 $x_0\in X$，使得 $T(x_0)=x_0$．称满足(4.1.2)式的映射 T 为**压缩（compression）映射**．

证 首先注意到，(4.1.2)式蕴含映射 T 的连续性．

事实上，$\forall x\in X$，$\forall\varepsilon>0$，$\exists\delta>0$，s. t. $\rho(x,y)<\delta$ 时，$\rho(T(x),T(y))<\varepsilon$（若 $\theta=0$，取 $\delta=\varepsilon$；若 $0<\theta<1$，取 $\delta=\dfrac{\varepsilon}{\theta}$）；由 $x\in X$ 的任意性，知 $T:X\to X$ 连续．

在 X 中任取一点 $x_1\in X$，令 $x_2=T(x_1)$，$x_3=T(x_2)$，$\cdots$，$x_{n+1}=T(x_n)$，$\cdots$．我们证明，这样得到的序列$\{x_n\}\subset X$ 是基本序列．事实上，

$$\begin{aligned}
&\rho(x_1,x_2)=\rho(T(x_0),T(x_1))\leqslant\theta\rho(x_0,x_1)=\theta\rho(x_0,T(x_0)),\\
&\rho(x_2,x_3)=\rho(T(x_1),T(x_2))\leqslant\theta\rho(x_1,x_2)=\theta^2\rho(x_0,T(x_0)),\\
&\vdots\\
&\rho(x_n,x_{n+1})=\rho(T(x_{n-1}),T(x_n))\leqslant\theta\rho(x_{n-1},x_n)=\theta\rho(T(x_{n-2}),T(x_{n-1}))\\
&\qquad\leqslant\theta^2\rho(x_{n-2},x_{n-1})\leqslant\cdots\leqslant\theta^n\rho(x_0,T(x_0)),\quad n=1,2,\cdots.
\end{aligned}$$

于是，$\forall p\in\mathbb{Z}^+$，当 $n\to+\infty$时，有

$$\begin{aligned}
\rho(x_n,x_{n+p})&\leqslant\rho(x_n,x_{n+1})+\rho(x_{n+1},x_{n+2})+\cdots+\rho(x_{n+p-1},x_{n+p})\\
&\leqslant\{\theta^n+\theta^{n+1}+\cdots+\theta^{n+p-1}\}\rho(x_0,T(x_0))=\frac{\theta^n(1-\theta^p)}{1-\theta}\rho(x_0,T(x_0))
\end{aligned}$$

$$\leqslant \frac{\theta^n}{1-\theta}\rho(x_0,T(x_0)) \to 0, \quad (因 0\leqslant\theta<1,故 \theta^n\to 0,n\to+\infty)$$

由此,序列$\{x_n\}$是X中的基本序列,再由X的完备性,存在$x_0\in X$,使得$\lim\limits_{n\to+\infty}x_n=x_0$.

在式$x_{n+1}=T(x_n)$两边取极限,并由映射T的连续性,得

$$x_0=\lim_{n\to+\infty}x_{n+1}=\lim_{n\to+\infty}T(x_n)=T(\lim_{n\to+\infty}x_n)=T(x_0),$$

故$x_0=T(x_0)$,因此x_0是映射T的不动点.

最后,证明不动点的惟一性.若映射T有两个不动点x_0,y_0,则

$$\rho(x_0,y_0)=\rho(T(x_0),T(y_0))\leqslant\theta\rho(x_0,y_0),$$

但是$0\leqslant\theta<1$,故$\rho(x_0,y_0)=0$,从而$x_0=y_0$,惟一性得证.

下面给出不动点定理的应用.

例 4.1.6 积分方程

$$x(t)=f(t)+\lambda\int_a^b K(t,s)x(s)\mathrm{d}s \tag{4.1.3}$$

称为**弗雷德霍姆(Fredholm)积分方程**,其中$f\in L^2([a,b])$给定,λ为参数,$K(x,s)(x,s\in[a,b])$为满足$\int_a^b\int_a^b|K(t,s)|^2\mathrm{d}t\mathrm{d}s<+\infty$的已知二元函数,称为**核(kernel)函数**,$x(t)$为所求的$L^2([a,b])$中的未知函数.则Fredholm积分方程对于充分小的$|\lambda|$,存在惟一的解$x=x(t)\in L^2([a,b])$.

证 为利用不动点定理4.1.5,定义映射$T:L^2([a,b])\to L^2([a,b])$,

$$(T(x))(t)\equiv T(x(t))=f(t)+\lambda\int_a^b K(t,s)x(s)\mathrm{d}s.$$

由Minkowski不等式,得估计式

$$\begin{aligned}\int_a^b\left|\int_a^b K(t,s)x(s)\mathrm{d}s\right|^2\mathrm{d}t&\leqslant\int_a^b\left[\int_a^b|K(t,s)|^2\mathrm{d}s\int_a^b|x(s)|^2\mathrm{d}s\right]\mathrm{d}t\\&=\int_a^b\int_a^b|K(t,s)|^2\mathrm{d}s\mathrm{d}t\int_a^b|x(s)|^2\mathrm{d}s<+\infty,\end{aligned}$$

故映射T将$x(t)$从$L^2([a,b])$映到$L^2([a,b])$.

当$|\lambda|$充分小时,有

$$\theta\equiv|\lambda|\left\{\int_a^b\int_a^b|K(t,s)|^2\mathrm{d}s\mathrm{d}t\right\}^{\frac{1}{2}}<1.$$

于是,再由Hölder不等式得

$$\begin{aligned}\rho(T(x),T(y))&=\left\{\int_a^b\left|\left\{f(t)+\lambda\int_a^b K(t,s)x(s)\mathrm{d}s\right\}-\left\{f(t)+\lambda\int_a^b K(t,s)y(s)\mathrm{d}s\right\}\right|^2\right\}^{\frac{1}{2}}\\&=|\lambda|\left\{\int_a^b\left|\int_a^b K(t,s)(x(s)-y(s))\mathrm{d}s\right|^2\mathrm{d}t\right\}^{\frac{1}{2}}\\&\leqslant|\lambda|\left\{\int_a^b\left(\int_a^b|K(t,s)|^2\mathrm{d}t\right)\mathrm{d}s\right\}^{\frac{1}{2}}\cdot\left\{\int_a^b|x(s)-y(s)|^2\mathrm{d}s\right\}^{\frac{1}{2}}\end{aligned}$$

$$= |\lambda|\left\{\int_a^b\left(\int_a^b|K(t,s)|^2\mathrm{d}t\right)\mathrm{d}s\right\}^{\frac{1}{2}}\rho(x,y) = \theta\rho(x,y),$$

因此,T 是压缩映射,由定理 4.1.5,存在惟一的函数 $x(t)\in L^2([a,b])$,满足 Fredholm 积分方程(4.1.3).

例 4.1.7 对于完备度量空间(X,ρ)到自身的映射 $T:X\to X$,如果存在常数 $\theta,0\leqslant\theta<1$,以及 $n_0\in\mathbb{Z}^+$,使得

$$\rho(T^{n_0}x,T^{n_0}y)\leqslant\theta\rho(x,y),\quad \forall(x,y)\in X\times X,$$

其中 T^n 是映射 T 的第 n 次迭代,$T^n(x)=T(T^{n-1}(x))$,则 T 在 X 中存在惟一的不动点.

证 映射 T 的第 n_0 次迭代 T^{n_0} 满足的条件 $\rho(T^{n_0}x,T^{n_0}y)\leqslant\theta\rho(x,y)$,即(4.1.2)式.根据不动点定理 4.1.5,映射 T^{n_0} 存在 X 中惟一的不动点 $x_0\in X$. 这个 x_0 也是映射 T 的不动点,因为

$$T^{n_0}(x_0)=x_0\Rightarrow T^{n_0}(T(x_0))=T^{n_0+1}(x_0)=T(T^{n_0}(x_0))=T(x_0)$$

$\Rightarrow$ $T(x_0)$是 T^{n_0} 的不动点,由 T^{n_0} 的不动点惟一性 $\Rightarrow T(x_0)=x_0\Rightarrow x_0$ 是 T 的不动点. 进而,x_0 也是 T 惟一的不动点,因为,若另有一个不动点 $x_1,T(x_1)=x_1$,则 $T(x_1)=x_1\Rightarrow$

$$T^{n_0}(x_1)=T^{n_0}(T(x_1))=T^{n_0+1}(x_1)=\underbrace{T\cdots T}_{n_0+1}(x_1)=\underbrace{T\cdots T}_{n_0}(T(x_1))=\underbrace{T\cdots T}_{n_0}(x_1)$$

$$=\underbrace{T\cdots T}_{n_0-1}(T(x_1))=\cdots=T(x_1)=x_1$$

$\Rightarrow$ x_1 是 T^{n_0} 的不动点 $\Rightarrow x_0=x_1$. 结论得证.

例 4.1.8 积分方程

$$x(t)=f(t)+\lambda\int_a^t K(t,s)x(s)\mathrm{d}s \tag{4.1.4}$$

称为**沃尔泰拉(Volterra)积分方程**,其中 $f\in C([a,b])$给定,λ 为参数;$K(x,s)$为定义在 $a\leqslant x\leqslant b,a\leqslant s\leqslant t$ 中的已知二元连续函数,称为**核函数**,$x(t)$为所求的 $C([a,b])$中的未知函数. 则 Volterra 积分方程对任意常数 λ 与任意 $f\in C([a,b])$,存在惟一的解 $x=x(t)\in C([a,b])$.

证 作映射 $T:C([a,b])\to C([a,b])$,

$$(T(x))(t)=f(t)+\lambda\int_a^t K(t,s)x(s)\mathrm{d}s, \tag{4.1.5}$$

则对于任意 $x_1(t),x_2(t)\in C([a,b])$,有

$$\begin{aligned}|(T(x_1))(t)-(T(x_2))(t)| &= |\lambda|\left|\int_a^t K(t,s)[x_1(s)-x_2(s)]\mathrm{d}s\right| \\ &\leqslant |\lambda|M(t-a)\max_{t\in[a,b]}|x_1(t)-x_2(t)| \\ &= |\lambda|M(t-a)\rho(x_1,x_2),\end{aligned} \tag{4.1.6}$$

其中 $M=\max\{|K(t,s)|:s\in[a,t],t\in[a,b]\}$，$\rho(x_1,x_2)=\max\limits_{t\in[a,b]}|x_1(t)-x_2(t)|$ 为 x_1 与 x_2 在 $C([a,b])$ 中的距离.

我们证明，(4.1.5)式中的映射 T 满足例4.1.7的条件 $\rho(T^{n_0}x,T^{n_0}y)\leqslant\theta\rho(x,y)$.

当 $n=1$ 时，(4.1.6)式给出 $|(T(x_1))(t)-(T(x_2))(t)|\leqslant|\lambda|M(t-a)\rho(x_1,x_2)$；

假设 $n=k$ 时，有 $|(T^k(x_1))(t)-(T^k(x_2))(t)|\leqslant|\lambda|^kM^k\dfrac{(t-a)^k}{k!}\rho(x_1,x_2)$，于是，

$$
\begin{aligned}
|(T^{k+1}(x_1))(t)-(T^{k+1}(x_2))(t)| &= |\lambda|\left|\int_a^t K(t,s)[T^kx_1(s)-T^kx_2(s)]\mathrm{d}s\right| \\
&\leqslant |\lambda|^{k+1}M^{k+1}\frac{1}{k!}\left\{\int_a^t(s-a)^k\mathrm{d}s\right\}\rho(x_1,x_2) \\
&= |\lambda|^{k+1}M^{k+1}\frac{(t-a)^{k+1}}{(k+1)!}\rho(x_1,x_2),
\end{aligned}
$$

根据归纳法，(4.1.5)式中的映射 T 对于一切 $n\in\mathbb{Z}^+$ 成立

$$
|(T^n(x_1))(t)-(T^n(x_2))(t)|\leqslant|\lambda|^nM^n\frac{(t-a)^n}{n!}\rho(x_1,x_2).
$$

这样，对于任意常数 λ，选取足够大的 $n\in\mathbb{Z}^+$，使得 $\theta=|\lambda|^nM^n\dfrac{(t-a)^n}{n!}<1$，故由例4.1.7中的结果，存在惟一的函数 $x(t)\in C([a,b])$，满足 Volterra 积分方程(4.1.4).

回顾在高等数学教程中学到的微分方程初始值问题 $\begin{cases}\dfrac{\mathrm{d}y}{\mathrm{d}x}=f(x,y),\\ y|_{x=x_0}=y_0,\end{cases}$ 在 $f(x,y)$ 满足关于 y 的 Lipschitz 条件 $|f(x,y)-f(x,y')|\leqslant M|y-y'|$ 下，证明了通过点 (x_0,y_0) 有惟一的一条积分曲线. 所用方法是将上述问题化为变动上限的 Volterra 积分方程 $y(x)=y_0+\displaystyle\int_{x_0}^x f(t,y(t))\mathrm{d}t$，并利用的逐次逼近法. 现在，只要定义映射

$$
(T(y))(x)\equiv T(y(x))=y_0+\int_{x_0}^x f(t,y(t))\mathrm{d}t, \tag{4.1.7}
$$

并且由

$$
\begin{aligned}
\rho(T(y_1),T(y_2)) &= \max_{|x-x_0|\leqslant\delta}\left|\int_{x_0}^x[f(t,y_1(t))-f(t,y_2(t))]\mathrm{d}t\right| \\
&\leqslant \max_{|x-x_0|\leqslant\delta}\left|\int_{x_0}^x M|y_1(t)-y_2(t)|\mathrm{d}t\right| \\
&\leqslant M\delta\max_{|x-x_0|\leqslant\delta}|y_1(t)-y_2(t)| \\
&= M\delta\rho(y_1,y_2),
\end{aligned}
$$

取 $\delta<1$，使得 $K\delta<1$，由(4.1.4)式中的 T 为 $C([x_0-\delta,x_0+\delta])$ 到自身的映射，利用例4.1.8中的结果，得到通过 (x_0,y_0) 的惟一的积分曲线.

在线性代数中,也遇到类似的问题:若方阵 $[a_{ij}]_{n\times n}\in\mathfrak{M}_n$ 满足 $\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n}|a_{ij}|^2<1$,则方程组

$$\xi_i-\sum_{j=1}^{n}a_{ij}\xi_j=b_i,\quad i=1,2,\cdots,n$$

有惟一解 $x^0=[x_i^0]_{n\times 1}\in\mathfrak{M}_1$,其中 $x=[\xi_i]_{n\times 1}$,$b=[b_i]_{n\times 1}\in\mathfrak{M}_1$.

事实上,设 $T:\mathfrak{M}_1\to\mathfrak{M}_1$ 是完备度量空间 $\mathfrak{M}_1(\overset{\text{iso}}{\longleftrightarrow}\mathbb{R}^n)$ 到自身的线性映射,满足

$$(Tx)_i=\sum_{j=1}^{n}a_{ij}\xi_j+b_i,\quad i=1,2,\cdots,n.$$

由于

$$\begin{aligned}\rho(Tx,Ty)&=\left\{\sum_{i=1}^{n}\left[\sum_{j=1}^{n}\left|a_{ij}(\xi_j^{(1)}-\xi_j^{(2)})\right|\right]^2\right\}^{\frac{1}{2}}\leqslant\left\{\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2\sum_{j=1}^{n}|\xi_j^{(1)}-\xi_j^{(2)}|^2\right\}^{\frac{1}{2}}\\&\leqslant\left\{\sum_{i=1}^{n}\sum_{j=1}^{n}|a_{ij}|^2\right\}^{\frac{1}{2}}\rho(x_1,x_2),\quad x_1,x_2\in\mathfrak{M}_1,\end{aligned}$$

令 $\theta=\sum\limits_{i=1}^{n}\sum\limits_{j=1}^{n}|a_{ij}|^2<1$,则映射 $T:\mathfrak{M}_1\to\mathfrak{M}_1$ 为压缩映射,因此存在惟一解 $x^0=[x_i^0]_{n\times 1}\in\mathfrak{M}_1$.

4.1.2 度量空间中的紧性

在点集拓扑基本知识一章中介绍了拓扑空间(X,τ)中各种紧性:紧致性、序列紧性、列紧性、可数紧性、局部紧致性,这些紧性都只与空间的拓扑有关. 当拓扑空间(X,τ)是度量空间(X,ρ)时,前面所证明的有关紧性的性质也成立,而在度量空间中,紧性还有更为丰富的性质. 此外,度量空间中可以定义的"全有界集"概念,与紧性有密切关系.

1. 度量空间中的全有界集

度量空间(X,ρ)中子集 $A\subset X$ 的有界性定义,已由定义 3.4.9 给出,而且定理 3.4.16 与定理 3.4.17 给出了有界性与紧致性的关系. 为研究度量空间的紧性,我们再引入全有界的概念.

定义 4.1.6(全有界性) 设(X,ρ)是度量空间,称子集 $A\subseteq X$ 为全有界的,若 A 具有有限 ε 网;亦即,对于任意给定的 $\varepsilon>0$,都存在相应的有限子集 $B=\{x_1,x_2,\cdots,x_n\}\subset X$,使得对任何 $x\in A$,存在 $x_j\in B=\{x_1,x_2,\cdots,x_n\}$,成立 $\rho(x,x_j)<\varepsilon$. 称 B 为 A 的**有限 ε 网(finite ε-net)**

显然,全有界性是有界性的推广(参看图 4.1.1 与图 4.1.2).

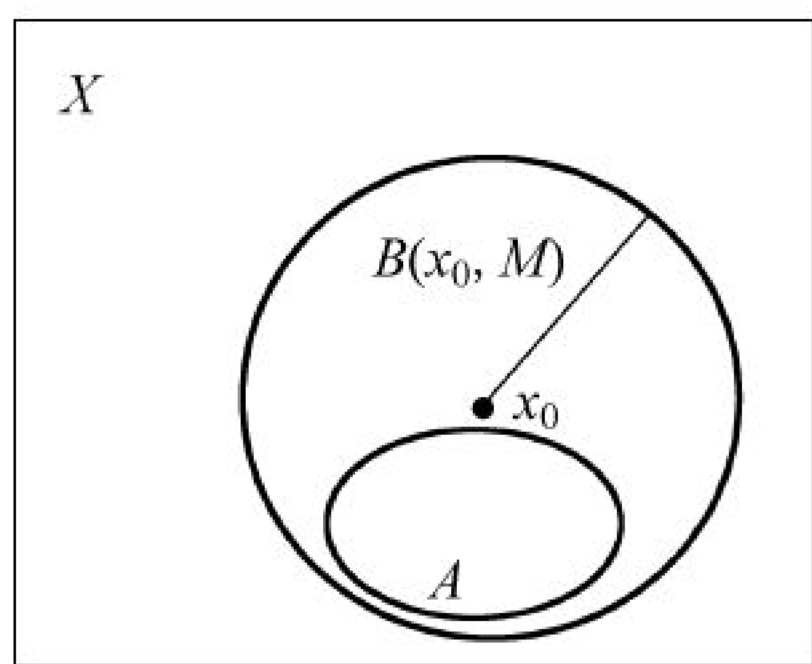


图 4.1.1 有界集 A

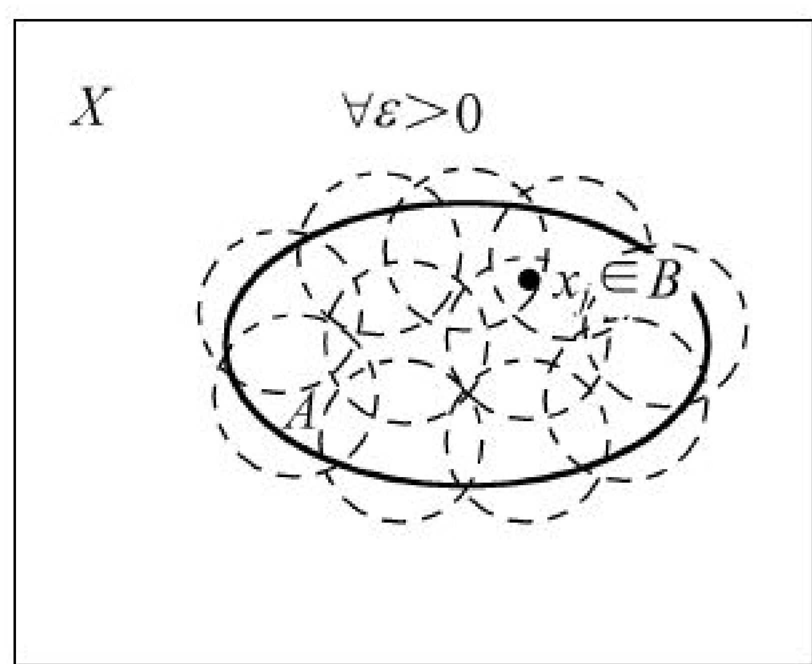


图 4.1.2 有限 ε 网

2. 度量空间中的各种紧性

1) 各种紧性的定义

定义 4.1.7 度量空间 (X,ρ) 中的各种紧性定义如下：

(1) **紧致性(compactness)** X 的任一开覆盖都存在有限子覆盖.(从"Heine-Borel 有限覆盖定理 —— 闭区间 $[a,b]$ 的任一个开覆盖必定存在有限子覆盖"得到启发.)

(2) **可数紧性(countable compactness)** X 的任一可数开覆盖都存在有限子覆盖.(从"闭区间套定理 —— 若 $[a_1,b_1]\supset\cdots\supset[a_n,b_n]\supset\cdots$，且 $\lim\limits_{n\to+\infty}(b_n-a_n)=0$，则 $\exists!\ x_0\in\bigcap\limits_{n=1}^{+\infty}[a_n,b_n]$"得到启发.)

(3) **序列紧性(sequentially compactness)** X 中任一无限序列都存在收敛到 X 中的元的子序列.(从"序列式 Bolzano-Weierstrass 定理—— 闭区间 $[a,b]$ 中任一无限序列都存在收敛到该区间中的元的子序列"得到启发.)

(4) **列紧性** X 中的任一无限子集都存在属于 X 的聚点.(从"子集式 Bolzano-Weierstrass 定理 ——闭区间 $[a,b]$ 中任一无限子集必存在属于 $[a,b]$ 的聚点" 得到启发.)

(5) **局部紧性(locally compactness)** X 中的每一点都存在紧致邻域.(从"欧氏空间 $\mathbb{R}^n$ 中每一点都有开球邻域，其闭包是紧致邻域"得到启发.)

定理 4.1.6 设 (X,ρ) 是度量空间，则有以下关系：

$$X\text{ 紧致}\Leftrightarrow X\text{ 可数紧}\Leftrightarrow X\text{ 序列紧}\Leftrightarrow X\text{ 列紧}\Rightarrow X\text{ 全有界}.$$

若 (X,ρ) 是完备度量空间，则

$$X\text{ 紧致}\Leftrightarrow X\text{ 可数紧}\Leftrightarrow X\text{ 序列紧}\Leftrightarrow X\text{ 列紧}\Leftrightarrow X\text{ 全有界}.$$

我们省略这个定理的证明，只强调一下：在度量空间 (X,ρ) 中，X 的紧致性、可数紧性、序列紧性、列紧性是彼此等价的.

这里再次强调"**连续映射保持紧致性**". 设 (X,ρ_1)，(Y,ρ_2) 是度量空间，$f:X\to Y$ 是连续映射，则紧致集 $A\subset X$ 的像集 $f(A)\subset Y$ 也是紧致集(定理 3.4.11). 因此，两个度量空间之间的**连续映射保持紧致性、可数紧性、序列紧性、列紧性.**

2）紧致性与完备性的关系

定理 4.1.7 设(X,ρ)为紧致度量空间，则 X 是全有界的完备空间.

证 设(X,ρ)为紧致度量空间，先证其全有界性. 任给 $\varepsilon>0$，则 $\forall x\in X$ 为心、ε 为半径的开球 $B(x,\varepsilon)=\{y\in X:\rho(x,y)<\varepsilon\}$所成的集合$\{B(x,\varepsilon):\forall x\in X,\forall\varepsilon>0\}$是 X 的一个开覆盖，$\bigcup_{x\in X}B(x,\varepsilon)=X$，故由 X 的紧致性，$\exists\{B(x_j,\varepsilon_j)\}_{j=1}^m$ 为有限开覆盖，因此$\{x_1,x_2,\cdots,x_m\}$就是 X 的有限 $\varepsilon=\max_j\{\varepsilon_1,\varepsilon_2,\cdots,\varepsilon_m\}$网，从而 X 是全有界的.

其次证明 X 是完备的. ① 证 X 的紧致性蕴含序列紧性. 设$\{x_n\}$是 X 的任一无限序列，设 $F_n=\overline{\{x_n,x_{n+1},\cdots,x_{n+p},\cdots\}}$，则$\{F_n\}$是 X 中的单调递减闭集序列，且满足有限交性质(集列中任意有限多个闭集的交非空). 假设 $\bigcap_{n=1}^{+\infty}F_n=\varnothing\Rightarrow U_n=X\setminus F_n$ 为开集，且 $\bigcup_{n=1}^{+\infty}U_n\supset X\left(因 X=\complement\left(\bigcap_{n=1}^{+\infty}F_n\right)=\bigcup_{n=1}^{+\infty}\complement F_n=\bigcup_{n=1}^{+\infty}U_n\right)\Rightarrow$ 由 X 的紧致性，$\exists\{U_{n_1},U_{n_2},\cdots,U_{n_m}\}$，s. t. $\bigcup_{k=1}^{m}U_{n_k}\supset X\Rightarrow\bigcap_{k=1}^{m}F_{n_k}=\bigcap_{k=1}^{m}\complement U_{n_k}=\complement\left(\bigcup_{k=1}^{m}U_{n_k}\right)=\varnothing\Rightarrow$ 与 $\bigcap_{k=1}^{m}F_{n_k}\neq\varnothing$ 矛盾 $\Rightarrow\bigcap_{n=1}^{+\infty}F_n\neq\varnothing$ $\Rightarrow\exists x\in\bigcap_{n=1}^{+\infty}F_n\neq\varnothing$，且 x 是$\{x_n\}$的聚点 $\Rightarrow\exists\{x_{n_j}\}$收敛到 $x\in X\Rightarrow X$ 中任意无限序列$\{x_n\}$都有收敛的子序列，亦即，X 的紧致性蕴含 X 的序列紧性.

② 证 X 的完备性. 任取 Cauchy 序列 $\{x_n\}\subset X$，据前面 ① 的证明，$\{x_n\}$必有聚点 $x\in F_1=\overline{\{x_n\}}$. 我们证明，此 Cauchy 序列的聚点就是它的极限 $\lim_{n\to+\infty}\rho(x_n,x)=0$：因若 $x\in X$ 是$\{x_n\}$的聚点，由 Cauchy 序列定义，$\forall\varepsilon>0$，$\exists N\in\mathbb{N}$，s. t. 当 $n,m>N$ 时，有 $\rho(x_n,x_m)<\varepsilon$；再由聚点的定义，$\forall\varepsilon>0$，$\exists p\geqslant N$，s. t. $\rho(x_p,x)<\varepsilon$；故当 $n>N$ 时，有 $\rho(x_n,x)\leqslant\rho(x_n,x_p)+\rho(x_p,x)<\varepsilon+\varepsilon=2\varepsilon$，此即 $\lim_{n\to+\infty}\rho(x_n,x)=0$，从而 X 是完备空间，定理得证.

3. 度量空间中子集的各种紧性

定义 4.1.8(子集的紧性) 设(X,ρ)是度量空间，$A\subset X$ 是子集.

(1) A 称为紧致的，若 A 的每个开覆盖都存在有限子覆盖；

(2) A 称为可数紧的，若 A 的每个可数开覆盖都存在有限子覆盖；

(3) A 称为序列紧的，若 A 中每个序列都含有收敛的子序列；

A 称为自序列紧的，若 A 中每个序列都含有收敛到 A 中某一点的子序列；

(4) A 称为列紧的，若 A 中每个无限子集都存在聚点；

A 称为自列紧的，若 A 中每个无限子集都存在属于 A 的聚点.

注 对于度量空间(X,ρ)本身，序列紧性与自序列紧性是不加区别的，因为收敛子序列的极限就在 X 中；列紧性与自列紧性也是不加区别的，因为无限子集的聚点也在 X 中. 但

对于子集 $A\subset X$ 就应当加以区别，因此要分为"序列紧"、"自序列紧"，也要区别"列紧"、"自列紧".

定理 4.1.8 (1) 设 (X,ρ) 是度量空间，$A\subset X$ 是 X 的子集，则有以下关系：

A 紧致 $\Leftrightarrow$ A 可数紧 $\Leftrightarrow$ A 自序列紧 $\Leftrightarrow$ A 自列紧 $\Rightarrow$ A 全有界；

(2) 设 (X,ρ) 是完备度量空间，$A\subset X$ 是 X 的子集，则有以下关系：

A 紧致 $\Leftrightarrow$ A 可数紧 $\Leftrightarrow$ A 自序列紧 $\Leftrightarrow$ A 自列紧 $\Leftrightarrow$ A 全有界；

(3) 在 $\mathbb{R}^n$ 中，$A\subset\mathbb{R}^n$ 是 $\mathbb{R}^2$ 的子集，则有如下关系：

A 紧致 $\Leftrightarrow$ A 可数紧 $\Leftrightarrow$ A 自序列紧 $\Leftrightarrow$ A 自列紧 $\Leftrightarrow$ A 有界闭.

例 4.1.9 实数集 $X=\mathbb{R}$ 是非紧致集，因为点列 $\{n\}$ 不含任何收敛子序列；进而，若 $X=\mathbb{R}$，$F=[a,b]$，$-\infty<a<b<+\infty$，则 F 为紧致集，因而也是自列紧集；$O=(a,b)$ 为列紧集，但非自列紧集.

例 4.1.10 设 $X=L^2[-\pi,\pi]$ 中的三角函数系 $\left\{\frac{1}{\sqrt{2\pi}},\frac{\cos t}{\sqrt{\pi}},\frac{\sin t}{\sqrt{\pi}},\cdots,\frac{\cos nt}{\sqrt{\pi}},\frac{\sin nt}{\sqrt{\pi}},\cdots\right\}$ 是有界的，但其中任何两个元的距离都等于 $\sqrt{2}$，故不存在收敛的子序列，因此此集不是列紧集合.

4. 赋范线性空间中的序列紧性判别法

在不同的赋范线性空间 $(X,\|x\|_X)$ 中，判断序列紧性有不同的定理.

1) $\mathbb{R}^n$ 中集合的序列紧性

定理 4.1.9 在 $\mathbb{R}^n$ 中，任何集 A 为序列紧集的充要条件是：A 是 $\mathbb{R}^n$ 中的有界闭集.

2) $C([a,b])$ 中集合的序列紧性

定理 4.1.10 (Arzela-Ascoli) 任何子集 $A\subset C([a,b])$ 为序列紧集的充要条件是：

(1) A 是**一致有界**的，亦即，存在常数 $M>0$，使得

$$\forall f\in A\Rightarrow |f(x)|\leqslant M,\quad \forall x\in[a,b],$$

也就是，按照 $C([a,b])$ 的范数 $\|f\|_{C([a,b])}=\max\limits_{x\in[a,b]}|f(x)|$，$A$ 是一致有界的；

(2) A 是**等度连续**的，亦即，$\forall\varepsilon>0$，$\exists\delta=\delta(\varepsilon)>0$，使得

$$\forall x_1,x_2\in[a,b],|x_1-x_2|<\delta\Rightarrow|f(x_1)-f(x_2)|<\varepsilon,\quad \forall f\in A.$$

3) $L^p([a,b])$ 中集合的序列紧性

定理 4.1.11 任何子集 $A\subset L^p([a,b])$ $(1<p<+\infty)$ 为序列紧集的充要条件是：

(1) A 是**一致有界**的，亦即，存在常数 $M>0$，使得

$$\forall f\in A\Rightarrow\int_{[a,b]}|f(x)|^p\mathrm{d}x\leqslant M,$$

也就是，按照 $L^p([a,b])$ 的范数 $\|f\|_{L^p([a,b])}=\left(\int_{[a,b]}|f(x)|^p\mathrm{d}x\right)^{\frac{1}{p}}$，$A$ 是一致有界的；

(2) A 是 L^p **一致平均连续**的，亦即，$\forall \varepsilon > 0, \exists \delta = \delta(\varepsilon) > 0$，使得

$$\forall h, 0 < h < \delta,\ \forall f \in A \Rightarrow \left\{\int_{[a,b]} |f_h(x) - f(x)|^p \mathrm{d}x\right\}^{\frac{1}{p}} < \varepsilon,$$

其中 f_h 定义如下：$\forall h>0, \forall x \in [a,b], f_h(x) = \dfrac{1}{2h}\displaystyle\int_{[x-h,x+h]} f(t)\mathrm{d}t$，并约定当 $x \notin [a,b]$ 时，$f(x) = 0$.

定理 4.1.10 与定理 4.1.11 的证明，可参看[6].

4）赋范线性空间的局部紧致性

定理 4.1.12 赋范线性空间$(X, \|x\|_X)$是局部紧致的，当且仅当 X 中的有界闭子集是紧致子集.

4.1.3 Banach 空间的基

第 2 章中介绍了线性空间$(X,+,\alpha\cdot)$的基 $S\subset X$，也就是 X 中的元由线性无关集 S 中的元"生成". 由定义知，X 需要有代数运算结构才能考虑"基". 因此本节考虑赋范线性空间$(X,+,\alpha\cdot, \|x\|_X)$中的基；又因任一度量空间都可以完备化，因此不失一般性，假设所考虑的$(X,+,\alpha\cdot, \|x\|_X)$是完备赋范线性空间，亦即 Banach 空间.

1. 常用的 Banach 空间

(1) $(\mathbb{R}^n, \|x\|_{\mathbb{R}^n})$.

$$\mathbb{R}^n = \{x = (x_1, x_2, \cdots, x_n): x_j \in \mathbb{R}, j = 1, 2, \cdots, n\}, \quad \|x\|_{\mathbb{R}^n} = \sqrt{\sum_{j=1}^{n} |x_j|^2}.$$

(2) $(l^p, \|x\|_{l^p}), 1 \leqslant p < +\infty$.

$$l^p = \left\{x = (x_1, x_2, \cdots): \sum_{j=1}^{+\infty} |x_j|^p < +\infty\right\}, \quad \|x\|_{l^p} = \left(\sum_{j=1}^{\infty} |x_j|^p\right)^{\frac{1}{p}}.$$

(3) $(l^\infty, \|x\|_{l^\infty})$.

$$l^\infty = \{x = (x_1, x_2, \cdots): |x_j| < +\infty, j \in \mathbb{Z}^+\}, \quad \|x\|_{l^\infty} = \sup_{j\in\mathbb{Z}^+} |x_j|.$$

(4) $(C([a,b]), \|f\|_{C([a,b])})$.

$$C([a,b]) = \{f: f \text{ 是}[a,b]\text{ 上的连续函数}\}, \quad \|f\|_{C([a,b])} = \max_{x\in[a,b]} |f(x)|.$$

(5) $(C^k([a,b]), \|f\|_{C^k([a,b])}), k \in \mathbb{N}$.

$$C^k([a,b]) = \{f: f^{(k)} \in C([a,b])\}, \quad \text{约定 } C^0([a,b]) = C([a,b]),$$

$$\|f\|_{C^k([a,b])} = \sum_{j=0}^{k} \max_{x\in[a,b]} |f^{(j)}(x)|.$$

(6) $(L^p([a,b]),\ \|f\|_{L^p([a,b])}),1\leqslant p<+\infty$.

$$L^p([a,b])=\left\{f:\int_{[a,b]}|f(x)|^p\mathrm{d}x<+\infty\right\},\quad \|f\|_{L^p([a,b])}=\left(\int_{[a,b]}|f(x)|^p\mathrm{d}x\right)^{\frac{1}{p}}.$$

(7) $(L^\infty([a,b]),\ \|f\|_{L^\infty([a,b])})$.

$$L^\infty([a,b])=\left\{f:\operatorname*{esssup}_{x\in[a,b]}|f(x)|<+\infty\right\},\quad \|f\|_{L^\infty[a,b]}=\operatorname*{esssup}_{x\in[a,b]}|f(x)|.$$

2. 有限维 Banach 空间

定义 4.1.9(Banach 空间的基) 设$(X,\|x\|_X)$是 Banach 空间,若存在 $n(n\geqslant 1)$个元 $e_1,e_2,\cdots,e_n\in X$,使得任一 $x\in X$ 可惟一地表示为

$$x=\sum_{j=1}^{n}\xi_j e_j,\quad \xi_j\in\mathbb{F},\quad j=1,2,\cdots,n,$$

则称$\{e_1,e_2,\cdots,e_n\}\subset X$ 为 X 的基,称 $\xi_1,\xi_2,\cdots,\xi_n\in\mathbb{F}$ 为 $x\in X$ 关于基$\{e_1,e_2,\cdots,e_n\}$的坐标;n 称为 X 的维数,X 称为(有限)n 维 Banach 空间,记为 $\dim X=n$.

若 X 只含有零元 $X=\{0\}$,则**约定** X 为零维 Banach 空间,$\dim\{0\}=0$.

于是,当 $n\in\mathbb{N}=\{0,1,\cdots\}$为有限非负整数时,$X$ 称为 n 维 Banach 空间,而集$\{e_1,e_2,\cdots,e_n\}\subset X$ 为 X 的基. 当 Banach 空间$(X,\|x\|_X)$不是有限维的,则称其为**无限维 Banach 空间**.

定义 4.1.10(拓扑同构) 设$(X,\|x\|_X)$,$(Y,\|y\|_Y)$为两个 Banach 空间,若满足

(1) 存在同构映射 $T:X\to Y$(一对一、满射、保持运算),使得 X 与 Y 为同构的线性空间;

(2) $T:X\to Y$ 是双方连续映射,亦即 T,T^{-1}为连续单射,

则称 X **与** Y **是拓扑同构的(topological isomorphic)**.

注 等距同构是拓扑同构的特例.

n 维 Banach 空间的拓扑同构定理叙述如下.

定理 4.1.13 设 n 为有限正整数,$n\in\mathbb{Z}^+$,则

(1) 任一个 n 维实 Banach 空间$(X,\|x\|_X)$与$\mathbb{R}^n$ 拓扑同构;任一个 n 维复 Banach 空间$(X,\|x\|_X)$与$\mathbb{C}^n$ 拓扑同构;因此,任一个 Banach 空间的有限维子空间是一个 Banach 空间,从而是一个闭子空间.

(2) 赋范线性空间$(X,\|x\|_X)$是有限维的,当且仅当$(X,\|x\|_X)$是局部紧的.

证 我们仅叙述此定理的证明思路,并就赋范线性空间加以叙述.

对于(1)，设$(X, \|x\|_X)$为 n 维赋范线性空间，取基$\{e_1, e_2, \cdots, e_n\}$，则$\forall x \in X$，$x = \sum_{j=1}^{n} \xi_j e_j$，$\xi_j \in \mathbb{F}$. 将$(\xi_1, \xi_2, \cdots, \xi_n)$视为$\mathbb{F}^n$ 中的点，作 X 到$\mathbb{F}^n$ 的映射 T：$T(x) = (\xi_1, \xi_2, \cdots, \xi_n)$，需要证明映射 T 的以下性质：

① $T: X \to \mathbb{F}^n$，$T(x) = (\xi_1, \xi_2, \cdots, \xi_n)$是 X 到$\mathbb{F}^n$ 的线性空间之间的同构映射；

② 对于基$\{e_1, e_2, \cdots, e_n\}$，存在常数 $c_1, c_2 > 0$，使对任意 $x = \sum_{j=1}^{n} \xi_j e_j \in X$，成立不等式

$$c_1 \|x\|_X \leqslant \left(\sum_{j=1}^{n} |\xi_j|^2\right)^{\frac{1}{2}} \leqslant c_2 \|x\|_X;$$

并由此得到 $c_1 \|x - y\|_X \leqslant \|T(x) - T(y)\|_{\mathbb{F}^n} \leqslant c_2 \|x - y\|_X$.

③ 利用上述不等式，即可证明 T 与 T^{-1}的连续性. 取$\{x_m\} \subset X$，在 X 中收敛于 $x_0 \in X$，亦即按 X 的范数有 $\|x_m - x_0\|_X \to 0$；故由 $\|T(x_m) - T(x_0)\|_{\mathbb{F}^n} \leqslant c_2 \|x_m - x_0\|_X$得 T 的连续性；再由 $T: X \to F^n$ 是线性空间之间的同构，从而可由 $c_1 \|T^{-1}(x) - T^{-1}(y)\|_{\mathbb{F}^n} \leqslant \|x - y\|_X$得到 T^{-1}的连续性.

对于(2)，需要用到重要的 **Riesz 引理**：设 X_0 是赋范线性空间 X 的真闭子空间，$X_0 \subsetneqq X$，则$\forall \varepsilon > 0$，$\exists x_0 \in X$，$\|x_0\|_X = 1$，s. t. $\|x - x_0\|_X \geqslant 1 - \varepsilon$ 对一切 $x \in X_0$ 成立，如图 4.1.3 所示.

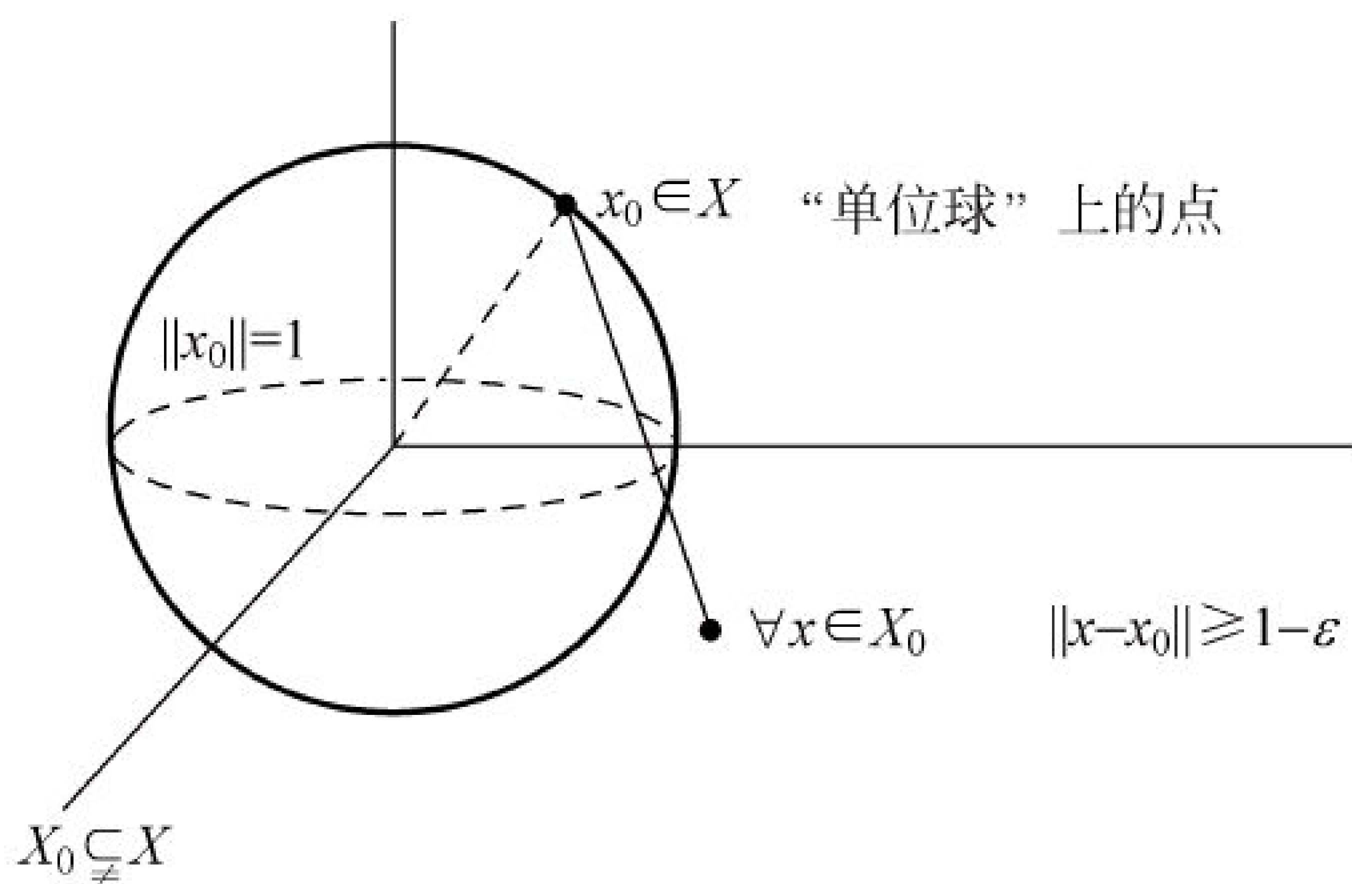


图 4.1.3 Riesz 引理示意

Riesz 引理的证明思路：因 $X_0 \subsetneqq X$，故$\exists x_1 \in X - X_0$，令 $d = \inf_{x \in X_0} \|x_1 - x\|_X$；其次因 $X_0 \subsetneqq X$ 为闭子空间，故 $d = \inf_{x \in X_0} \|x_1 - x\|_X > 0$(度量空间中闭集与集外一点之间距离必大于零)；再由下确界的性质，任取 $0 < \varepsilon < 1$，对于$\frac{d}{1-\varepsilon} > d$，存在点 $x_1' \in X_0$，使得 $\|x_1 - x_1'\|_X < \frac{d}{1-\varepsilon}$.

令 $x_0 = \frac{x_1 - x_1'}{\|x_1 - x_1'\|_X}$，得 $\|x_0\|_X = 1$. 这样，$\forall x \in X_0$，有

$$\|x - x_0\|_X = \left\|x - \frac{x_1 - x_1'}{\|x_1 - x_1'\|_X}\right\|_X = \frac{1}{\|x_1 - x_1'\|_X}\Big\| \|x_1 - x_1'\|_X x - x_1 + x_1' \Big\|_X$$

$$= \frac{1}{\|x_1 - x_1'\|_X} \left\| (\|x_1 - x_1'\|_X x + x_1') - x_1 \right\|_X, \tag{4.1.8}$$

由于 $x_1' \in X_0$，$x \in X_0$，而 X_0 为线性子空间，故 $\|x_1 - x_1'\|_X x + x_1' \in X_0$，因此

$$\left\| (\|x_1 - x_1'\|_X x + x_1') - x_1 \right\|_X \geqslant d,$$

且由(4.1.8)式，有

$$\begin{aligned} \|x - x_0\|_X &= \frac{1}{\|x_1 - x_1'\|_X} \left\| (\|x_1 - x_1'\|_X x + x_1') - x_1 \right\|_X \\ &\geqslant \frac{d}{\|x_1 - x_1'\|_X} > 1 - \varepsilon, \end{aligned}$$

由此 Riesz 引理得证.

为证(2)，在必要性的证明中，利用(1)的结果，$(X, \|x\|_X)$ 与 $\mathbb{F}^n$ 拓扑同构，因此，同构映射将 X 中的有界闭集映到 $\mathbb{F}^n$ 中的有界闭集，反之亦然. 而 $\mathbb{F}^n$ 中的有界闭集 A，当且仅当 A 是紧致集.

于是，$(X, \|x\|_X)$ 的每一点的开球邻域的闭包是有界闭集，从而是紧致集. 故 $(X, \|x\|_X)$ 是局部紧赋范线性空间.

为证充分性，采用反证法. 设 $(X, \|x\|_X)$ 是无限维的局部紧赋范线性空间. 令 $S = \{x \in X: \|x\|_X = 1\}$ 为 $(X, \|x\|_X)$ 中的单位球面，可视为有界闭集，从而是紧致集(本章习题 14). 任取 $x_1 \in S$，记 X_1 是由 x_1 张成的线性子空间，显然是有限维的；由于 X_1 与 $\mathbb{R}^1$ 拓扑同构，因此 X_1 是 X 的有限维真闭子空间，故由 Riesz 引理，$\exists x_2 \in S$，使得 $\|x - x_2\|_X \geqslant \frac{1}{2}$，$\forall x \in X_1$. 因 $x_1 \in X_1$，也有 $\|x_1 - x_2\|_X \geqslant \frac{1}{2}$.

由 x_1, x_2 张成一个线性子空间，记为 X_2，则 X_2 也是 X 的有限维线性真子空间，同理，也是闭的，于是再由 Riesz 引理，$\exists x_3 \in S$，使得 $\|x - x_3\|_X \geqslant \frac{1}{2}$，$\forall x \in X_2$. 特别地，因 x_1，$x_2 \in X_2$，故也有 $\|x_3 - x_1\|_X \geqslant \frac{1}{2}$，$\|x_3 - x_2\|_X \geqslant \frac{1}{2}$. 继续此过程，得到序列 $x_1, x_2, \cdots, x_k, \cdots \in S$，满足 $\|x_k - x_l\|_X \geqslant \frac{1}{2}$，$k \neq l$，$k, l \geqslant 1$.

但因序列 $\{x_k\} \subset S$ 满足 $\|x_k - x_l\|_X \geqslant \frac{1}{2}$，$k \neq l$，$k, l \geqslant 1$，故没有收敛的子序列. 这与 S 的紧致性矛盾. 充分性得证.

3. 无限维 Banach 空间

定义 4.1.11(Schauder 基) 设 $(X, \|x\|_X)$ 是 Banach 空间，称**无穷序列** $\{e_1, e_2, \cdots, e_n, \cdots\} = \{e_n\} \subset X$ 为 X 的 **Schauder 基**，若 $\forall x \in X$，可惟一地表示为

$$x=\sum_{n=1}^{+\infty}c_ne_n,\quad c_n\in\mathbb{F},\ n\in\mathbb{Z}^+,$$

这里级数是按照 Banach 空间$(X,\|x\|_X)$范数收敛的，亦即，$\lim\limits_{n\to+\infty}\left\|x-\sum\limits_{k=1}^{n}c_ke_k\right\|_X=0$. 此时，Schauder 基$\{e_n\}\subset X$也简称为$X$的基，$\{c_n\}\subset\mathbb{F}$称为$x\in X$ **关于基$\{e_n\}$的坐标**；Schauder 基是**可数基**，具有 Schauder 基的 Banach 空间X是无穷维空间.

例 4.1.11 $(l^p,\|x\|_{l^p})$，$1\leqslant p<+\infty$.

对于 $l^p=\left\{x=(x_1,x_2,\cdots):\sum\limits_{j=1}^{+\infty}|x_j|^p<+\infty\right\}$，令 $\|x\|_{l^p}=\left(\sum\limits_{j=1}^{\infty}|x_j|^p\right)^{\frac{1}{p}}$，则$(l^p,\|x\|_{l^p})$是一个无穷维 Banach 空间. 证明如下.

(1) l^p 是数域$\mathbb{F}=\mathbb{C}$或$\mathbb{R}$上的线性空间. 关于线性运算封闭，由 Minkowski 不等式

$$\left(\sum_{j=1}^{\infty}|x_j+y_j|^p\right)^{\frac{1}{p}}\leqslant\left(\sum_{j=1}^{\infty}|x_j|^p\right)^{\frac{1}{p}}+\left(\sum_{j=1}^{\infty}|y_j|^p\right)^{\frac{1}{p}}$$

得到；

(2) $\|x\|_{l^p}=\left(\sum\limits_{j=1}^{\infty}|x_j|^p\right)^{\frac{1}{p}}$ 是一个范数. 关于范数的第一、第二条件 $\|x\|_{l^p}\geqslant0$，$x=0\Leftrightarrow\|x\|_{l^p}=0$，$\|\alpha x\|_{l^p}=|\alpha|\|x\|_{l^p}$显然. 而第三个条件 $\|x+y\|_{l^p}\leqslant\|x\|_{l^p}+\|y\|_{l^p}$也立即由 Minkowski 不等式得到；

(3) l^p 是一个 Banach 空间. 在范数为 $\|x\|_{l^p}=\left(\sum\limits_{j=1}^{\infty}|x_j|^p\right)^{\frac{1}{p}}$ 之下，l^p 是完备赋范线性空间. 在距离 $\rho(x,y)=\left(\sum\limits_{j=1}^{\infty}|x_j-y_j|^p\right)^{\frac{1}{p}}$ 之下，完备性由空间中的 Cauchy 序列收敛到空间中的点($\mathbb{R}$ 的连续性)得到；

(4) l^p 的 Schauder 基为

$$e_1=(1,0,0,\cdots,0,\cdots),\quad e_2=(0,1,0,\cdots,0,\cdots),\cdots,\quad e_n=(\underbrace{0,\cdots,0,1}_{n},0,\cdots),\cdots.$$

对于任一点 $x=(x_1,x_2,\cdots,x_j,\cdots)\in l^p$，有 $x=\sum\limits_{j=1}^{+\infty}x_je_j$，且 $\lim\limits_{k\to+\infty}\left\|x-\sum\limits_{j=1}^{k}x_je_j\right\|_{l^p}=0$.

例 4.1.12 $(L^p([a,b]),\|f\|_p)$，$1\leqslant p<+\infty$.

对于 $L^p([a,b])=\left\{f:\int_{[a,b]}|f(x)|^p\mathrm{d}x<+\infty\right\}$，令 $\|f\|_p=\left(\int_{[a,b]}|f(x)|^p\mathrm{d}x\right)^{\frac{1}{p}}$，则$(L^p([a,b]),\|f\|_p)$是一个无穷维 Banach 空间.

(1) $L^p([a,b])$是数域$\mathbb{F}=\mathbb{C}$或$\mathbb{R}$上线性空间. 线性运算的封闭性，由 Minkowski 不等式

$$\left(\int_{[a,b]}|f(x)+g(x)|^p\mathrm{d}x\right)^{\frac{1}{p}}\leqslant\left(\int_{[a,b]}|f(x)|^p\mathrm{d}x\right)^{\frac{1}{p}}+\left(\int_{[a,b]}|g(x)|^p\mathrm{d}x\right)^{\frac{1}{p}}$$

得到;

(2) $\|f\|_p=\left(\int_{[a,b]}|f(x)|^p\mathrm{d}x\right)^{\frac{1}{p}}$ 是一个范数. 关于范数的第一条件:非负性 $\|f\|_p\geqslant 0$ 显然;

条件 $\|f\|_p=0\Leftrightarrow f\sim 0$ 是指 $\|f\|_p=0\Leftrightarrow f=0$, a. e. (表示 f 几乎处处为零,亦即 $\forall x\in[a,b]\backslash E$ 蕴含 $f(x)=0$,其中 E 为 Lebesgue 可测集,且 $m(E)=0$).

先证"$f\sim 0\Rightarrow\|f\|_p=0$". 由 $\int_{[a,b]}|f(x)|^p\mathrm{d}x=\int_{[a,b]\backslash E}|f(x)|^p\mathrm{d}x+\int_E|f(x)|^p\mathrm{d}x=0$,第一个积分为零,因为在$[a,b]\backslash E$ 上 $f(x)=0$,第二个积分为零,因为 $m(E)=0$.

反之,"$\|f\|_p=0\Rightarrow f\sim 0$"由以下推理得到:设 $n\in\mathbb{Z}^+$,则 $\int_{[a,b]}|f(x)|^p\mathrm{d}x\geqslant\int_{E(|f|^p\geqslant\frac{1}{n})}|f(x)|^p\mathrm{d}x\geqslant\frac{1}{n}m\left(E\left(|f|^p\geqslant\frac{1}{n}\right)\right)$,故 $\|f\|_p=0\Rightarrow m\left(E\left(|f|^p\geqslant\frac{1}{n}\right)\right)=0$. 进而,由 $E(f\neq 0)=\bigcup_{n=1}^{+\infty}E\left(|f|^p\geqslant\frac{1}{n}\right)\Rightarrow m(E(f\neq 0))\leqslant\sum_{n=1}^{+\infty}m\left(E\left(|f|^p\geqslant\frac{1}{n}\right)\right)=0$ 得到 $f\sim 0$.

关于范数的第二条件 $\|\alpha f\|_p=|\alpha|\|f\|_p$ 显然;

关于范数的第三条件 $\|f+g\|_p\leqslant\|f\|_p+\|g\|_p$ 也是由 Minkowski 不等式得到.

(3) $L^p([a,b])$ 是 Banach 空间. 范数为 $\|f\|_p=\left(\int_{[a,b]}|f(x)|^p\mathrm{d}m\right)^{\frac{1}{p}}$,距离为 $\rho(f,g)=\left(\int_{[a,b]}|f(x)-g(x)|^p\mathrm{d}m\right)^{\frac{1}{p}}$. 完备性证明已经在例 4.1.5 中给出.

(4) $L^p([a,b])(1\leqslant p<+\infty)$是无穷维的 Banach 空间. 由 $C([a,b])\subset L^p([a,b])\subset L^1([a,b])$,我们证明 $C([a,b])$是无穷维的. 事实上,$\forall n\in\mathbb{Z}^+$,在 $C([a,b])$中都存在 n 个线性无关的函数组,例如$\{x^k:k=0,1,\cdots,n,x\in[a,b]\}\subset C([a,b])$. 如果 $C([a,b])$是有限维 Banach 空间,例如 n 维空间,则 $C([a,b])$中任一组 $n+1$ 个函数的函数组都是线性相关的. 但这不可能. 因此,$C([a,b])$与 $L^p([a,b])(p\geqslant 1)$都是无穷维的 Banach 空间.

4.1.4 Hilbert 空间的直交系与直交展开

$\mathbb{R}^n(n\in\mathbb{Z}^+)$,$l^p(1\leqslant p\leqslant+\infty)$是点集所成的空间,$C^k([a,b])(k\in\mathbb{N})$,$L^p([a,b])(1\leqslant p\leqslant+\infty)$是函数空间. 函数空间大多是无穷维的,故考虑其 Schauder 基. 函数空间中的元的序列,称为函数系. 将每个函数用函数系中的元来表示,称为"按照函数系展开". 完备内积空间(Hilbert 空间)中函数的直交展开有较为成熟的理论,并且也最有用.

1. 内积空间中的直交性

设$(X,(x,y))$为内积空间，(x,y)为其内积(定义 2.1.6、定义 2.1.7). 首先，简单回顾$(X,(x,y))$的常用性质.

(1) 内积的线性与共轭线性

若$(X,(x,y))$是实内积空间，则内积(x,y)关于x,y都是线性的；若$(X,(x,y))$是复内积空间，则(x,y)关于x是线性的、关于y是共轭线性的.

(2) 内积空间是一个赋范线性空间

若$(X,(x,y))$是内积空间，令$\|x\|_X=\sqrt{(x,x)}$，则$(X,\|x\|_X)$是一个赋范线性空间，并且成立 Schwarz 不等式 $|(x,y)|\leqslant\|x\|_X\|y\|_X$；

(3) 内积空间$(X,(x,y))$的内积(x,y)是x与y的连续函数

若$\{x_n\}\subset X$，$\{y_n\}\subset X$，且$\lim\limits_{n\to+\infty}\|x_n-x\|_X=0$，$\lim\limits_{n\to+\infty}\|y_n-y\|_X=0$，$x\in X$，$y\in Y$，这里$\|z\|_X=\sqrt{(z,z)}$是内积空间$X$中的元$z\in X$的范数. 则$\lim\limits_{n\to+\infty}|(x_n,y_n)-(x,y)|=0$.

事实上，有

$$\begin{aligned}|(x_n,y_n)-(x,y)|&\leqslant|(x_n,y_n)-(x,y_n)|+|(x,y_n)-(x,y)|\\&=|(x_n-x,y_n)|+|(x,y_n-y)|\\&\leqslant\|x_n-x\|_X\|y_n\|_X+\|x\|_X\|y_n-y\|_X\\&\leqslant M_1\|x_n-x\|_X+\|x\|_X\|y_n-y\|_X\to 0,\quad n\to+\infty,\end{aligned}$$

这里用到了X中收敛序列$\{y_n\}$的有界性.

(4) 内积空间$(X,(x,y))$中的极化恒等式

若$(X,(x,y))$是实内积空间，由内积导出的范数$\|x\|_X=\sqrt{(x,x)}$满足两个恒等式

$$(x,y)=\frac{1}{4}(\|x+y\|_X^2-\|x-y\|_X^2),$$

$$\|x+y\|_X^2+\|x-y\|_X^2=2(\|x\|_X^2+\|y\|_X^2);$$

后者称为**内积空间中的平行四边形公式**；

若$(X,(x,y))$是复内积空间，则有恒等式

$$(x,y)=\frac{1}{4}(\|x+y\|_X^2-\|x-y\|_X^2+\mathrm{i}\|x+\mathrm{i}y\|_X^2-\mathrm{i}\|x-\mathrm{i}y\|_X^2).$$

(5) 内积空间$(X,(x,y))$中的“勾股定理”

若子集$\mathfrak{A}=\{x_1,x_2,\cdots,x_k\}\subset X$中的元都是两两相互直交的，记$x=x_1+x_2+\cdots+x_k$，则对于范数$\|x\|_X=\sqrt{(x,x)}$，有**推广勾股定理**$\|x\|_X^2=\|x_1\|_X^2+\|x_2\|_X^2+\cdots+\|x_k\|_X^2$.

欧氏空间中直交的概念可以推广到 Hilbert 空间中.

定义 4.1.12(直交性) 设$(X,(x,y))$为 Hilbert 空间，对于非零元$x,y\in X$，若$x,y\in X\Rightarrow(x,y)=0$，则称$x,y$互相**直交(orthogonal)**，记为$x\perp y$；也称为正交.

对于元 $x\in X$ 与子集 $M\subset X$，若 x 与 M 中的每个元 $y\in M$ 直交，则称 x **与 M 直交**，记为 $x\perp M$. 与子集 $M\subset X$ 中的每个元 $y\in M$ 直交的元 $x\in X$ 的全体，记为

$$M^{\perp}=\{x\in X: x\perp y, \forall y\in M\}.$$

定义 4.1.13（直交分解） 设 $(X,(x,y))$ 为 Hilbert 空间，$Y\subset X$ 是 X 闭子空间，则对任一元 $x\in X$，存在惟一的**直交分解（orthogonal decomposition）**

$$x=y+z,\quad y\in Y,\quad z\in Y^{\perp},$$

称 $y\in Y$ 为 $x\in X$ **在闭子空间 Y 中的直交投影（orthogonal projection）**.

定理 4.1.14 设 $(X,(x,y))$ 是 Hilbert 空间，M 是 X 的闭子空间，则 $\forall x\in X$，存在惟一的直交分解

$$x=x_0+y,\quad \exists x_0\in M,\ y\in M^{\perp}.$$

证 不失一般性，设 $M\underset{\neq}{\subset} X$ 为真子空间. 若 $x\notin M$，则 $\alpha=\inf\limits_{y\in M}\|x-y\|_X>0$. 于是，$\exists\{x_n\}\subset M$, s. t. $\alpha=\lim\limits_{n\to+\infty}\|x-x_n\|_X \Rightarrow \dfrac{x_n+x_m}{2}\in M$ 蕴含 $\left\|x-\dfrac{x_n+x_m}{2}\right\|_X\geqslant\alpha \Rightarrow$

$$\begin{aligned}\|x_m-x_n\|_X^2&=\|x_m-x-x_n+x\|_X^2\\&=2\|x_m-x\|_X^2+2\|x_n-x\|_X^2-4\left\|x-\frac{x_m+x_n}{2}\right\|_X^2\\&\leqslant 2\|x_m-x\|_X^2+2\|x_n-x\|_X^2-4\alpha^2\to 2\alpha^2+2\alpha^2-4\alpha^2=0\end{aligned}$$

$\Rightarrow \lim\limits_{n,m\to+\infty}\|x_m-x_n\|_X=0 \Rightarrow M$ 闭，故完备 $\Rightarrow \exists x_0\in M$, s. t. $\lim\limits_{n\to+\infty}\|x_n-x_0\|_X=0$，且对于 $x\in X\backslash M$, $\|x-x_0\|_X=\alpha$.

还需证明 $(x-x_0)\perp M$. 任取 $z\in M, z\neq 0$，对任意复数 $\lambda\in\mathbb{C}$，有 $x_0+\lambda z\in M$

$\Rightarrow \alpha^2\leqslant\|x-(x_0+\lambda z)\|_X^2=\|x-x_0\|_X^2-\bar{\lambda}(x-x_0,z)-\lambda(z,x-x_0)+|\lambda|^2\|z\|^2$

$\Rightarrow \bar{\lambda}(x-x_0,z)+\lambda(z,x-x_0)-|\lambda|^2\|z\|^2\leqslant 0$ （由上式移项，并由 $\|x-x_0\|_X=\alpha$ 得此不等式）

$\Rightarrow |(x-x_0,z)|^2+|(z,x-x_0)|^2-|(x-x_0,z)|^2\leqslant 0$ $\left(\text{取上式 } \lambda=\dfrac{(x-x_0,z)}{\|z\|^2}\text{，则得此不等式}\right)$

$\Rightarrow |(x-x_0,z)|^2\leqslant 0 \Rightarrow (x-x_0,z)=0 \Rightarrow$ 由 $z\in M$ 的任意性，得到 $(x-x_0)\perp M$

$\Rightarrow$ 令 $y=x-x_0$，得 $x=x_0+y$，其中 $x_0\in M, y\in M^{\perp}$. 定理得证.

定义 4.1.14（标准直交系） 设 $(X,(x,y))$ 为 Hilbert 空间，若 X 的子集

$$\{e_1,e_2,\cdots,e_n,\cdots\}\subset X$$

满足 $(e_i,e_j)=\begin{cases}1, & i=j,\\ 0, & i\neq j,\end{cases}$ 则称 $\{e_n\}$ 是 X 的一个**标准直交系（normal orthogonal system）**.

例 4.1.13 在例 4.1.11 中,取 $p=2$,得到 $l^2=\left\{x=(x_1,x_2,\cdots):\sum_{j=1}^{+\infty}|x_j|^2<+\infty\right\}$,内积定义为 $(x,y)=\sum_{j=1}^{+\infty}x_jy_j$,由 Schwarz 不等式 $\left(\sum_{j=1}^{\infty}|x_jy_j|\right)\leqslant\left(\sum_{j=1}^{\infty}|x_j|^2\right)^{\frac{1}{2}}\left(\sum_{j=1}^{\infty}|y_j|^2\right)^{\frac{1}{2}}$,知 $x,y\in l^2\Rightarrow|(x,y)|<+\infty$. 易证 $(x,y)=\sum_{j=1}^{+\infty}x_jy_j$ 满足内积的定义. 故 l^2 是 Hilbert 空间. l^2 的 Schauder 基 $\{e_n=(\underbrace{0,\cdots,0,1}_{n},0,\cdots):n\in\mathbb{Z}^+\}$ 就是 l^2 的一个标准直交系.

例 4.1.14 在例 4.1.12 中,取 $p=2$,得 $L^2([a,b])=\left\{f:\int_{[a,b]}|f(x)|^2\mathrm{d}x<+\infty\right\}$,内积定义为 $(f,g)=\int_{[a,b]}f(x)\overline{g(x)}\mathrm{d}x$. 利用 Schwarz 不等式 $\int_{[a,b]}|f(x)\overline{g(x)}|\mathrm{d}x\leqslant\left(\int_{[a,b]}|f(x)|^2\mathrm{d}x\right)^{\frac{1}{2}}\left(\int_{[a,b]}|g(x)|^2\mathrm{d}x\right)^{\frac{1}{2}}$ 得 $f,g\in L^2([a,b])\Leftrightarrow|(f,g)|<+\infty$, 且 $fg\in L^1([a,b])$. 进而,(f,g) 满足内积的定义,从而 $L^2([a,b])$ 是一个 Hilbert 空间.

函数族 $\left\{\frac{1}{\sqrt{2\pi}}\mathrm{e}^{\mathrm{i}nx}\right\}_{n\in\mathbb{Z}}$ 是 $L^2([-\pi,\pi])$ 中的一个标准直交系.

例 4.1.15 加权 Hilbert 空间 $L^2([a,b],\omega(x))$,$\omega(x)$ 是 $[a,b]$ 上的正值可测函数. 令

$$L^2([a,b],\omega(x))=\left\{f(x):\int_{[a,b]}\omega(x)|f(x)|^2\mathrm{d}x<+\infty\right\},$$

由广义 Schwarz 不等式

$$\int_{[a,b]}\omega(x)|f(x)\overline{g(x)}|\mathrm{d}x\leqslant\left(\int_{[a,b]}\omega(x)|f(x)|^2\mathrm{d}x\right)^{\frac{1}{2}}\cdot\left(\int_{[a,b]}\omega(x)|g(x)|^2\mathrm{d}x\right)^{\frac{1}{2}}$$

得到 $L^2([a,b],\omega(x))$ 对运算与加权内积的封闭性,并令 $(f,g)=\int_{[a,b]}\omega(x)f(x)\overline{g(x)}\mathrm{d}x$ 为**加权内积**,于是得到**加权 Hilbert** 空间 $L^2([a,b],\omega(x))$.

例 4.1.16 加权 Hilbert 空间 $L^2\left([-1,1],\frac{1}{\sqrt{1-x^2}}\right)$ 中,有一个称为第一类切比雪夫(Chebyshev)多项式的直交系

$$\{T_n(x)=\cos(n\arccos x):x\in[-1,1]\}_{n\in\mathbb{N}}.$$

令 $\widetilde{T}_n(x)=\begin{cases}\frac{1}{\sqrt{\pi}}T_n(x), & n=0,\\ \frac{2}{\sqrt{\pi}}T_n(x), & n\neq0,\end{cases}$ 则 $\{\widetilde{T}_n(x):x\in[-1,1]\}_{n\in\mathbb{N}}$ 是 $[-1,1]$ 上的标准直交系.

例 4.1.17 加权 Hilbert 空间 $L^2((-\infty,+\infty),\mathrm{e}^{-x^2})=L^2(\mathbb{R},\mathrm{e}^{-x^2})$

显然,$\{1,x,x^2,\cdots,x^n,\cdots\}\subset L^2(\mathbb{R},\mathrm{e}^{-x^2})$. 对于权函数 $\omega(x)=\mathrm{e}^{-x^2}$,有 $\omega'(x)=$

$-2xe^{-x^2}$，$\omega''(x)=(4x^2-2)e^{-x^2}$，…，$\omega^{(n)}(x)=y_n(x)e^{-x^2}$，其中 $y_n(x)$ 是 x 的 n 次多项式，最高次项的系数是 $(-2)^n$. 可以证明，$\{y_n(x)\}_{n\in\mathbb{N}}$ 是 $\{1,x,x^2,\cdots,x^n,\cdots\}$ 经过直交化的多项式族. 令 $H_n(x)=\dfrac{y_n(x)}{\|y_n\|_{L^2(\mathbb{R},e^{-x^2})}}$，则 $\{H_n(x)\}_{n\in\mathbb{N}}$ 是 $L^2(\mathbb{R},e^{-x^2})$ 中的标准完整直交系.

2. Hilbert 空间中的直交展开

在 Hilbert 空间 $(X,(x,y))$ 中定义了直交性后，可以得到 Hilbert 空间的许多重要性质.

1）Hilbert 空间中的 Fourier 展开

定义 4.1.15（Fourier 系数） 设 $(X,(x,y))$ 为 Hilbert 空间，$\{e_n\}\subset X$ 是 X 的标准直交系，对于任一 $x\in X$，称数列

$$c_n=(x,e_n),\quad n=1,2,\cdots \tag{4.1.9}$$

为 x 关于标准直交系 $\{e_n\}$ 的**第 n 个 Fourier 系数.**

定理 4.1.15 设 $(X,(x,y))$ 为 Hilbert 空间，$\{e_n\}\subset X$ 是 X 的标准直交系，则对于任一 $x\in X$，不等式

$$\sum_{n=1}^{+\infty}|(x,e_n)|^2\leqslant\|x\|_X^2 \tag{4.1.10}$$

成立.

不等式(4.1.10)称为 **Bessel 不等式**，这里 $(x,e_n)=c_n$ 是 x 的 Fourier 系数，$\|x\|_X=\sqrt{(x,x)}$ 是 Hilbert 空间作为 Banach 空间，元 x 的范数.

2）Hilbert 空间中直交系的完全性与完整性

定义 4.1.16（Parseval 等式） 设 $(X,(x,y))$ 为 Hilbert 空间，$\{e_n\}\subset X$ 是 X 的标准直交系. 对于任一 $x\in X$，等式

$$\sum_{n=1}^{+\infty}|(x,e_n)|^2=\|x\|_X^2 \tag{4.1.11}$$

称为 **Parseval 等式**，其中 $(x,e_n)=c_n$，$n=1,2,\cdots$，$\|x\|_X=\sqrt{(x,x)}$.

定义 4.1.17（完全性） 设 $(X,(x,y))$ 为 Hilbert 空间，$\{e_n\}\subset X$ 是 X 的标准直交系. 若对于任一 $x\in X$，都有 Parseval 等式成立，则称**标准直交系 $\{e_n\}$ 是完全的(completed).**

注 完全性表明了空间的能量守恒性，亦即 Parseval 等式揭示了能量守恒律.

定义 4.1.18（完整性） 设 $(X,(x,y))$ 为 Hilbert 空间，$\{e_n\}\subset X$ 是 X 的标准直交系. 若对于任一 $x\in X$，有

$$(x,e_n)=0,\quad n=1,2,\cdots\Rightarrow x=0, \tag{4.1.12}$$

则称**标准直交系 $\{e_n\}$ 具有完整性(integrity).**

注 完整性表明,凡与标准直交系$\{e_n\}$中任一元都直交的非零元$x\in X$,已全部包含在$\{e_n\}$中. 标准完整直交系$\{e_1,e_2,\cdots,e_n,\cdots\}$中包含了所有彼此直交的非零元.

定理 4.1.16 设$(X,(x,y))$为 Hilbert 空间,$\{e_n\}\subset X$是X的标准直交系,则下列 4 条等价:

(1) $\{e_n\}$是完全的;

(2) $\{e_n\}$是完整的;

(3) $\{e_n\}$所张成的子空间$L\subset X$在X中稠密;

(4) $\forall x\in X$,级数$\sum\limits_{n=1}^{+\infty}(x,e_n)e_n$在$(X,(x,y))$中收敛到$x$.

证 (1) ⇒ (3) 设$\{e_n\}$是完全系,即$\sum\limits_{n=1}^{+\infty}|(x,e_n)|^2=\|x\|_X^2$,这等价于$\forall x\in X$,成立$\left\|\sum\limits_{k=1}^{n}c_ke_k-x\right\|_X\to 0(n\to+\infty)$,其中$c_k=(x,e_k)$;故表示由$\{e_n\}$张成的子空间$L$在$X$中稠密.

(3) ⇒ (1) 任取$x\in X$,记$(x,e_k)=c_k(k=1,2,\cdots)$. 因$L$在$X$中稠密,故$\forall\varepsilon>0$,$\exists\{\alpha_k\}$, $\exists n>0$, s. t. 对$s_n=\sum\limits_{k=1}^{n}\alpha_ke_k$,成立$\|x-s_n\|_X<\varepsilon$.

我们证明,使得等式$\|x-s_n\|_X=\left\|x-\sum\limits_{k=1}^{n}\alpha_ke_k\right\|_X$取最小值,当且仅当$\alpha_k=c_k$. 事实上,由$\left(x-\sum\limits_{k=1}^{n}\alpha_ke_k,e_j\right)=(x,e_j)-\alpha_j=0(j=1,2,\cdots,n)$,故$x-\sum\limits_{k=1}^{n}\alpha_ke_k$与$e_j$直交,当且仅当$\alpha_j=c_j$. 进而,由内积空间的广义勾股定理,$\left\|x-\sum\limits_{k=1}^{n}\alpha_ke_k\right\|_X^2=\left\|x-\sum\limits_{k=1}^{n}c_ke_k+\sum\limits_{k=1}^{n}(c_k-\alpha_k)e_k\right\|_X^2=\left\|x-\sum\limits_{k=1}^{n}c_ke_k\right\|_X^2+\left\|\sum\limits_{k=1}^{n}(c_k-\alpha_k)e_k\right\|_X^2=\left\|x-\sum\limits_{k=1}^{n}c_ke_k\right\|_X^2+\sum\limits_{k=1}^{n}|c_k-\alpha_k|^2$.

于是,当且仅当$\alpha_k=c_k$时,式$\|x-s_n\|_X=\left\|x-\sum\limits_{k=1}^{n}\alpha_ke_k\right\|_X$达到最小值,故$\left\|x-\sum\limits_{k=1}^{n}c_ke_k\right\|_X<\varepsilon$. 再由 Parseval 等式的充要条件,知等式$\sum\limits_{n=1}^{+\infty}|(x,e_n)|^2=\|x\|_X^2$成立.

(1) ⇒ (2) 设$\{e_n\}$为完全系,证$\{e_n\}$是完整系:取$x\in X$适合$(x,e_n)=0(n=1,2,\cdots)$,由(1)的假设,$\{e_n\}$是完全系,Parseval 等式给出$\|x\|_X^2=\sum\limits_{n=1}^{+\infty}|(x,e_n)|^2=\sum\limits_{n=1}^{+\infty}|c_n|^2=0$,于是,$x=0$,因此完整性得证.

(2) ⇒ (1) 设$\{e_n\}$为完整系,证$\{e_n\}$是完全的,只需证完整系$\{e_n\}$张成的子空间L在X中稠密. 反证,若不然,即$\overline{L}\neq X$,则$\exists x\in X\backslash\overline{L}$,$x\neq 0$,故$x=x_0+y$, s. t. $x_0\in\overline{L}$,$y\perp\overline{L}$(直

交分解定理 4.1.14)，显然 $y\neq 0$. 但是，y 与所有 $\{e_n\}$ 直交，这与 $\{e_n\}$ 的完整性矛盾，故 $\overline{L}=X$，从而，由(3)与(1)等价性得到 $\{e_n\}$ 的完全性.

(4)与前三条的等价性显然.

定理 4.1.17 设$(X,(x,y))$是可分 Hilbert 空间，则必存在 X 的标准完整直交系 $\{e_n\}\subset X$.

此处省略本定理的具体证明过程，只讲述证明思路. 由空间的可分性，存在可数的稠密子集 $A=\{x_n\}\subset X$，使得 $X=\overline{A}$. 将 $A=\{x_n\}$ 中的元采用如下的 Schmidt 直交化方法，作出一个标准完整直交系 $B=\{e_n\}$，而内积空间$(X,(x,y))$正好与由 $B=\{e_n\}$ 张成的子空间相同.

Schmidt 直交化方法：设 $\{x_n\}\subset(X,(x,y))$为内积空间中的可数子集序列，取 $e_1=\dfrac{x_1}{\|x_1\|_X}$，则 $\|e_1\|_X=1$. 令 $y_2=x_2-(x_2,e_1)e_1$，则 $y_2\neq 0$，$(y_2,e_1)=0$，故 $y_2\perp e_1$；取 $e_2=\dfrac{y_1}{\|y_1\|_X}$，则 $\|e_2\|_X=1$，$e_2\perp e_1$. 令 $y_3=x_3-(x_3,e_1)e_1-(x_3,e_2)e_2$，则 $y_3\neq 0$，$(y_3,e_1)=0$，$(y_3,e_2)=0$，故 $y_3\perp e_1$，$y_3\perp e_2$；取 $e_3=\dfrac{y_2}{\|y_2\|_X}$，则 $\|e_3\|_X=1$，$e_3\perp e_1$，$e_3\perp e_2$. 继续以上步骤，令 $y_4=x_4-(x_4,e_1)e_1-\cdots$，便可得到标准完整直交系 $\{e_n\}$.

定理 4.1.18 对于无穷维 Hilbert 空间$(X,(x,y))$，有

(1) 任一个实可分 Hilbert 空间$(X,(x,y))$必定与实 l^2 空间等距同构；任一个复可分 Hilbert 空间$(X,(x,y))$必定与复 l^2 空间等距同构；

(2) 所有可分 Hilbert 空间$(X,(x,y))$都是等距同构的.

利用 Hilbert 空间$(X,(x,y))$的标准完整直交系 $\{e_n\}$，将 X 中的元 $x\in X$ 进行 Fourier 展开，于是其 Fourier 系数 $\{c_n\}$，$(c_n=(x,e_n))$与 l^2 空间之间的对应关系就成为$(X,(x,y))$与 l^2 的等距同构.

例 4.1.18 $[0,1]$区间上的 Rademarch 函数系 $\{\varphi_n(x)\}_{n\in\mathbb{N}}$，其中

$$\varphi_0(x)=\begin{cases}1, & x\in[0,2^{-1}),\\ -1, & x\in(2^{-1},1),\end{cases}\qquad \varphi_n(x)=\varphi_0(2^n x),\ n=1,2,\cdots$$

是$[0,1]$上的标准直交系，但它不是完整系；将其完整化，得到 Walsh 函数系 $\{\mathrm{wal}_n(x)\}_{n\in\mathbb{N}}$ 才是标准完整直交系：

$$\mathrm{wal}_0(x)=1,\quad \mathrm{wal}_k(x)=\prod_{j:k'_{-j}=1}\varphi_j(x),\quad k\in\mathbb{Z}^+,$$

其中 $k\in\mathbb{Z}^+$ 的二进表示为 $k=(k_{-N+1},k_{-N},\cdots,k_{-1},k_0)$，而 $k'_{-j}=k_{-j}\oplus k_{-(j+1)}$.

Hilbert 空间中的直交展开，特别是 Fourier 级数，在高等数学中有详细的介绍. 在自然科学的诸多分支中，如 Fourier 分析以及在 Fourier 分析基础上发展起来的信号分析、各种变换(包括 Laplace 变换、小波变换)等，都起着极为重要的作用.

4.2 算子理论

4.2.1 Banach 空间上的线性算子

空间,可以理解为研究对象的"运动场",自变量在一个"运动场"中运动,因变量在同一个或另一个"运动场"中运动. 所以,具有各种性质的空间以及各种空间之间的映射,就成为数学的主要研究对象. 本节研究 Banach 空间上的线性算子.

1. Banach 空间上的有界线性算子

1) 有界线性算子

定义 4.2.1(线性算子) 设$(X, \|x\|_X)$,$(Y, \|y\|_Y)$为数域$\mathbb{F}$上的 Banach 空间,$D\subseteq X$为 X 的子空间,若映射 $T:D\to Y$ 满足$\forall x_1, x_2, x\in D$,$\forall \alpha\in\mathbb{F}$,成立

$$T(x_1+x_2)=T(x_1)+T(x_2),\quad T(\alpha x)=\alpha T(x),$$

则称 T 是 X 到 Y 的**线性算子**,D 称为算子 T 的**定义域**.

当 $D=X$,$Y=\mathbb{R}$ 或$\mathbb{C}$时,称 T 为 X 上的**线性泛函**.

定义 4.2.2(有界线性算子) 设$(X, \|x\|_X)$,$(Y, \|y\|_Y)$为数域$\mathbb{F}$上的 Banach 空间,$D\subseteq X$ 为 X 的子空间,若映射 $T:D\to Y$ 是 X 到 Y 的线性算子,且存在常数 $M>0$,满足

$$\|T(x)\|_Y\leqslant M\|x\|_X,\quad \forall x\in D, \tag{4.2.1}$$

则称 T 是 X 到 Y 的**有界线性算子(bounded linear operator)**;若 $Y=\mathbb{F}$,且存在常数 $M>0$,满足

$$|T(x)|\leqslant M\|x\|_X,\quad \forall x\in D, \tag{4.2.2}$$

则称 T 是 X 上的**有界线性泛函(bounded linear functional)**.

定义 4.2.3(算子的范数) 设$(X, \|x\|_X)$、$(Y, \|y\|_Y)$为数域$\mathbb{F}$上的 Banach 空间,若映射 $T:D\subseteq X\to Y$ 是 X 到 Y 的有界线性算子,称满足不等式

$$\|T(x)\|_Y\leqslant M\|x\|_X,\quad \forall x\in D$$

的 $M>0$ 的下确界,为**算子 T 的范数(norm of operator T)**,记为$\|T\|$,亦即

$$\|T\| = \inf\{M: \|T(x)\|_Y\leqslant M\|x\|_X, \forall x\in D\}. \tag{4.2.3}$$

算子范数也可以等价地表示为

$$\|T\| = \sup\{\|T(x)\|_Y: \|x\|_X\leqslant 1, \forall x\in D\} \tag{4.2.4}$$

或

$$\|T\| = \sup\{\|T(x)\|_Y: \|x\|_X = 1, \forall x\in D\}. \tag{4.2.5}$$

例 4.2.1 决定$\mathbb{R}^n$到$\mathbb{R}^n$上的有界线性算子 $T:\mathbb{R}^n\to\mathbb{R}^n$ 的形式.

事实上,$T:\mathbb{R}^n\to\mathbb{R}^n$ 由一个 $n\times n$ 阶方阵 $A=[a_{ij}]_{1\leqslant i,j\leqslant n}\in\mathfrak{M}_n$决定. 元素

$$x=(\xi_1,\xi_2,\cdots,\xi_n)\in\mathbb{R}^n\to y=(\eta_1,\eta_2,\cdots,\eta_n)\in\mathbb{R}^n$$

通过变换 $\eta_i=\sum_{j=1}^n a_{ij}\xi_j(i=1,2,\cdots,n)$,给出. $T(x)=y$ 表示为

$$T(x)=[a_{ij}]_{n\times n}[x_i]_{n\times 1}=\begin{bmatrix}a_{11}&\cdots&a_{1n}\\ \vdots&\ddots&\vdots\\ a_{n1}&\cdots&a_{nn}\end{bmatrix}\begin{bmatrix}\xi_1\\ \vdots\\ \xi_n\end{bmatrix}=\begin{bmatrix}a_{11}\xi_1+\cdots+a_{1n}\xi_n\\ \vdots\\ a_{n1}\xi_1+\cdots+a_{nn}\xi_n\end{bmatrix}$$

$$=\Big[\sum_{j=1}^n a_{ij}\xi_j\Big]_{n\times 1}=\begin{bmatrix}\eta_1\\ \vdots\\ \eta_n\end{bmatrix}=y.$$

不难证明,$T(x)=y^{\mathrm{T}}=[\eta_1\cdots\eta_n]\left(\text{矩阵 } y=\begin{bmatrix}\eta_1\\ \vdots\\ \eta_n\end{bmatrix}\text{的转置}\right)$是 $\mathbb{R}^n\to\mathbb{R}^n$ 的有界线性算子,因为它的线性是易见的;有界性则由 Cauchy 不等式

$$\left(\sum_{i=1}^n\eta_i^2\right)^{\frac{1}{2}}=\left(\sum_{i=1}^n\left(\sum_{j=1}^n a_{ij}\xi_j\right)^2\right)^{\frac{1}{2}}\leqslant\left(\sum_{i=1}^n\sum_{j=1}^n a_{ij}^2\right)^{\frac{1}{2}}\left(\sum_{j=1}^n\xi_j^2\right)^{\frac{1}{2}}=\left(\sum_{i=1}^n\sum_{j=1}^n a_{ij}^2\right)^{\frac{1}{2}}\|x\|_{\mathbb{R}^n}$$

得到,并且有 $\|T\|\leqslant\left(\sum_{i=1}^n\sum_{j=1}^n a_{ij}^2\right)^{\frac{1}{2}}$.

例 4.2.2 设 $C([a,b])$为$[a,b]$上的连续函数的全体. $f\in C([a,b])$的范数定义为

$$\|f\|_{C([a,b])}=\max_{x\in[a,b]}|f(x)|,$$

使得 $C([a,b])$为 Banach 空间. 对于 $f\in C([a,b])$,定义积分算子 J 为

$$J(f)(g)=\int_a^b f(x)g(x)\mathrm{d}x,\quad\forall g\in C([a,b]),$$

则 $J:C([a,b])\to\mathbb{R}$ 是 $C([a,b])$上的有界线性泛函.

事实上,由

$$|J(f)(g)|=\left|\int_a^b f(x)g(x)\mathrm{d}x\right|\leqslant(b-a)\|f\|_{C([a,b])}\|g\|_{C([a,b])}$$

得到.

例 4.2.3 设 $C^k([a,b])$为$[a,b]$上的 k 次连续可微函数的全体. $f\in C^k([a,b])$的范数定义为

$$\|f\|=\sum_{j=0}^k\max_{x\in[a,b]}|f^{(j)}(x)|,$$

使得 $C^k([a,b])$为 Banach 空间. 定义**求导算子**

$$\frac{\mathrm{d}}{\mathrm{d}x}:C^k([a,b])\to C^{k-1}([a,b]),\quad k\geqslant 1,$$

则$\dfrac{\mathrm{d}}{\mathrm{d}x}$是$C^k([a,b])$上的线性算子.

但是,$\dfrac{\mathrm{d}}{\mathrm{d}x}$不是有界的,例如,在$C^1([0,1])$中,取$f_n(x)=\sin nx$,则$\|f_n(x)\|_{C([0,1])}=1$,却有

$$\left\|\frac{\mathrm{d}}{\mathrm{d}x}f_n(x)\right\|_{C([0,1])}=\|n\cos nx\|_{C([0,1])}=n\to+\infty.$$

所以,并非每个线性算子都是有界的.

在赋范线性空间中,有界线性算子是连续的,反之亦然.

定理 4.2.1 设$(X,\|x\|_X)$,$(Y,\|y\|_Y)$为数域$\mathbb{F}$上的Banach空间,则映射$T:X\to Y$是X到Y的有界线性算子,当且仅当T是X到Y的连续线性算子;也当且仅当T把X中的有界集映到Y中的有界集.

证 (1) 必要性 设$T:X\to Y$是有界线性算子,则存在常数$M>0$,满足$\|T(x)\|_Y\leqslant M\|x\|_X$,$\forall x\in X$ $\Rightarrow$ 对于任一$x\in X$,取$x_n\in X$,$n=1,2,\cdots$,s.t. $\|x_n-x\|_X\to 0(n\to+\infty)$,则$\|T(x_n-x)\|_Y\leqslant M\|x_n-x\|_X\Rightarrow\|Tx_n-Tx\|_Y=\|T(x_n-x)\|_Y\to 0(n\to+\infty)\Rightarrow T:X\to Y$连续(由连续性定义).

充分性 设$T:X\to Y$是连续线性算子. 用反证法,若T是无界的,则$\forall n\in\mathbb{N}$,$\exists x_n\in X$,$x_n\neq 0$,s.t. $\|T(x_n)\|_Y\geqslant n\|x_n\|_X\Rightarrow$ 令$y_n=\dfrac{x_n}{n\|x_n\|_X}$,则$\|y_n\|_Y=\left\|\dfrac{x_n}{n\|x_n\|_X}\right\|_Y=\dfrac{1}{n}\to 0\Rightarrow$ 由T的连续性,对于上述$y_n\in Y$,有$T(y_n)\to 0(n\to+\infty)\Rightarrow\|T(y_n)\|_Y\to 0(n\to+\infty)$.

另一方面,由T的线性,$T(y_n)=T\left(\dfrac{x_n}{n\|x_n\|_X}\right)=\dfrac{1}{n\|x_n\|_X}T(x_n)$,故

$$\|T(y_n)\|_Y=\left\|T\left(\frac{x_n}{n\|x_n\|_X}\right)\right\|_Y=\left\|\frac{1}{n\|x_n\|_X}T(x_n)\right\|_Y=\frac{1}{n\|x_n\|_X}\|T(x_n)\|_Y\geqslant 1,$$

这与$\|T(y_n)\|_Y\to 0(n\to+\infty)$矛盾. 于是,$T$是有界的.

(2) 必要性 设$T:X\to Y$是有界线性算子,则存在常数$M>0$,满足$\|T(x)\|_Y\leqslant M\|x\|_X$,$\forall x\in X\Rightarrow$ 取有界集$A\subset X$,则存在正数$K>0$,使得$\forall x\in A$蕴含$\|x\|_X\leqslant K\Rightarrow$ 联合上两式,得$\|T(x)\|_Y\leqslant M\|x\|_X\leqslant MK$,$\forall x\in A\Rightarrow$ 像集$T(A)$是Y中的有界子集.

充分性 设线性映射$T:X\to Y$把X中的有界集映到Y中的有界集$\Rightarrow S=\{x\in X:\|x\|_X=1\}$为$X$中的单位球面,显然,$S$是$X$中的有界集$\Rightarrow$ 根据假设,$T(S)\subset Y$是Y中的有界集,故存在正数$M>0$,使得$\forall x\in S=\{x\in X:\|x\|_X=1\}$蕴含$\|T(x)\|_Y\leqslant M\Rightarrow\forall x\in X$,$x\neq 0$蕴含$\dfrac{x}{\|x\|_X}\in S\Rightarrow\left\|T\left(\dfrac{x}{\|x\|_X}\right)\right\|_Y\leqslant M\Rightarrow\forall x\in X$蕴含$\|T(x)\|_Y\leqslant M\|x\|_X\Rightarrow T:X\to Y$是有界算子.

今后我们将有界线性算子与连续线性算子同时使用.

2）有界线性算子空间

定义 4.2.4(算子空间) 设$(X,\|x\|_X)$,$(Y,\|y\|_Y)$为数域$\mathbb{F}$上的赋范线性空间，记由X到Y的有界线性算子$T:X\to Y$的全体为

$$\mathfrak{B}(X,Y)=\{T:X\to Y,T\text{是有界线性算子}\},$$

它是X到Y的线性算子集合．在$\mathfrak{B}(X,Y)$中定义加法运算+与数乘运算$\alpha\cdot$为$(T+S)(x)=T(x)+S(x)$，$\forall T,S\in\mathfrak{B}(X,Y)$；$(\alpha T)(x)=\alpha T(x)$，$\forall T\in\mathfrak{B}(X,Y)$，$\alpha\in\mathbb{F}$，则$\mathfrak{B}(X,Y)$成为$\mathbb{F}$上的线性空间；对于每个算子$T\in\mathfrak{B}(X,Y)$，定义其范数(定义4.2.3)$\|T\|$，则$\mathfrak{B}(X,Y)$也成为赋范线性空间$(\mathfrak{B}(X,Y),\|T\|)$．我们称$(\mathfrak{B}(X,Y),\|T\|)$为**有界线性算子空间**，也称其为**连续线性算子空间**．

2. Banach空间上有界线性算子的重要性质

对于Banach空间$(X,\|x\|_X)$与$(Y,\|y\|_Y)$，有界线性算子$T\in\mathfrak{B}(X,Y)$有重要性质．我们介绍三个重要定理：**开映射定理、逆算子定理、闭图像定理**．

1）开映射定理

定理 4.2.2 设$T:X\to Y$是Banach空间X到Y的有界线性算子$T\in\mathfrak{B}(X,Y)$，且T为满射，$T(X)=Y$，则T把X中的开集映到Y中的开集，亦即，

$$T\in\mathfrak{B}(X,Y)\xrightarrow{T(X)=Y}T\text{ 是开映射}.$$

证 为证T是开映射，分为三步．

(1) 证明下述等价性：

$\forall$开集$W\subset X$，像集$T(W)\subset Y$为开集$\Leftrightarrow\exists\delta>0$, s. t. $T(B(0,1))\supset U(0,\delta)$，这里$B(0,1)$是$X$中以零元为心、1为半径的开球，$U(0,\delta)$是$Y$中以零元为心、$\delta$为半径的开球．

必要性 假设$\forall$开集$W\subset X$，像集$T(W)\subset Y$为开集$\Rightarrow$因$B(0,1)\subset X$是X中以0(Banach空间X的零元)为中心、1为半径的开球是开集，故据假设条件，像集$T(B(0,1))\subset Y$为Y中的开集$\Rightarrow$由T的线性，知$T(0)=0\in T(B(0,1))\Rightarrow$对于$Y$中开集$T(B(0,1))$，必存在$\delta>0$，使得$Y$中的开球$U(0,\delta)$含在开集$T(B(0,1))$中，即$T(B(0,1))\supset U(0,\delta)$．

充分性 假设$\exists\delta>0$, s. t. $T(B(0,1))\supset U(0,\delta)\Rightarrow\forall$开集$W\subset X$，像集$T(W)\subset Y$为开集$\Rightarrow$对开集$W\subset X$，由$T$为满射，得到：$\forall y_0\in T(W)$，$\exists x_0\in W$, s. t. $y_0=T(x_0)\Rightarrow\exists r>0$, s. t. $B(x_0,r)\subset W\Rightarrow y_0\in T(B(x_0,r))\subset T(W)$；

另一方面，因$T\in\mathfrak{B}(X,Y)$，T连续，$\exists\varepsilon>0$, s. t. 邻域$U(T(x_0),r\varepsilon)\subset T(B(x_0,r))\Rightarrow$令$\delta=r\varepsilon>0$，则$\exists\delta>0$, s. t. $U(y_0,\delta)\subset T(B(x_0,r))\subset T(W)\Rightarrow\forall y_0\in T(W)$，$\exists\delta>0$, s. t. $U(y_0,\delta)\subset T(W)\Rightarrow T(W)$是开集．故(1)得证．

(2) 证明 $\forall B(0,1)\subset X$，$\exists\delta>0$，s. t. $T(B(0,1))\supset U(0,\delta)$. 分两步进行证明.

第一步，证 $\exists\delta>0$，s. t. $\overline{T(B(0,1))}\supset U(0,3\delta)$.

因 T 为满射，故 $Y=T(X)=\bigcup_{n=1}^{+\infty}T(B(0,n))\Rightarrow$ 因 Y 为 Banach 空间，故至少存在一个 $n_0\in\mathbb{Z}^+$，s. t. $T(B(0,n_0))$具有非空内部，即 $T(B(0,n_0))$至少有一内点 $y_0\in T(B(0,n_0))\Rightarrow$ $\exists r>0$，s. t. 邻域 $U(y_0,r)\subset T(B(0,n_0))\subset\overline{T(B(0,n_0))}\Rightarrow$ 因 X 中的球 $B(0,n_0)$对称，故 $x\in B(0,n_0)$蕴含$(-x)\in B(0,n_0)\Rightarrow T(B(0,n_0))\subset Y$ 对称，亦即 $y=T(x)\in T(B(0,n_0))$蕴含 $(-y)=-T(x)=T(-x)\in T(B(0,n_0))\Rightarrow U(-y_0,r)\subset\overline{T(B(0,n_0))}\Rightarrow$ 成立下列包含关系：

$$U(0,r)\subset\frac{1}{2}U(y_0,r)+\frac{1}{2}U(-y_0,r)\subset\overline{T(B(0,n_0))}.$$

上式中第一个包含关系是因为

$$y\in U(0,r)\Rightarrow y=y_1+y_2,\text{其中 }y_1\in\frac{1}{2}U(y_0,r),y_2\in\frac{1}{2}U(-y_0,r)$$

$$\Rightarrow\|y-0\|_Y\leqslant\|y_1-y_0\|_Y+\|y_2-(-y_0)\|_Y\leqslant\frac{r}{2}+\frac{r}{2}=r;$$

第二个包含关系是因为

$$U(y_0,r)\subset\overline{T(B(0,n_0))}\Rightarrow\frac{1}{2}U(y_0,r)\subset\frac{1}{2}\overline{T(B(0,n_0))},$$

$$U(-y_0,r)\subset\overline{T(B(0,n_0))}\Rightarrow\frac{1}{2}U(-y_0,r)\subset\frac{1}{2}\overline{T(B(0,n_0))}.$$

故 $\frac{1}{2}U(y_0,r)+\frac{1}{2}U(-y_0,r)\subset\overline{T(B(0,n_0))}\Rightarrow$ 由 $T(\alpha x)=\alpha T(x)$，以及 $U(0,r)\subset\overline{T(B(0,n_0))}=\overline{T(n_0B(0,1))}=n_0\overline{T(B(0,1))}$得到

$$\frac{1}{n_0}U(0,r)\subset\overline{T(B(0,1))}\Rightarrow U\left(0,\frac{r}{n_0}\right)\subset\overline{T(B(0,1))},$$

取 $3\delta=\frac{r}{n_0}$，则 $\exists\delta>0$，s. t. $\overline{T(B(0,1))}\supset U(0,3\delta)$.

第二步，证对于上述 $\delta>0$，$T(B(0,1))\supset U(0,\delta)$. 也只需证

$$\forall y_0\in U(0,\delta),\exists x_0\in B(0,1),\text{ s. t. }T(x_0)=y_0.$$

亦即，只需

$$\forall y_0\in U(0,\delta),\text{求方程 }T(x)=y_0\text{ 在 }B(0,1)\text{ 中的解 }x_0.$$

用逐次逼近法. $\forall y_0\in U(0,\delta)$，$\exists x_1\in B\left(0,\frac{1}{3}\right)$，s. t. $\|y_0-T(x_1)\|_Y<\frac{\delta}{3}\Rightarrow$ 由第一步的结论，$\forall y_0\in U(0,\delta)\subset\overline{T\left(B\left(0,\frac{1}{3}\right)\right)}$，有两种情况：

$$U\left(y_0,\frac{\delta}{3}\right)\subset\overline{T\left(B\left(0,\frac{1}{3}\right)\right)},\exists x_1\in B\left(0,\frac{1}{3}\right),\text{ s. t. }\|y_0-T(x_1)\|_Y<\frac{\delta}{3};$$

或者

$$U\left(y_0,\frac{\delta}{3}\right)\not\subset \overline{T\left(B\left(0,\frac{1}{3}\right)\right)},\exists k>3,\text{ s. t. } U\left(y_0,\frac{\delta}{k}\right)\subset \overline{T\left(B\left(0,\frac{1}{3}\right)\right)}.$$

然而在第二种情况时，也有：$\exists x_1\in B\left(0,\frac{1}{3}\right)$, s. t. $\|y_0-T(x_1)\|_Y<\frac{\delta}{k}<\frac{\delta}{3}$. 因此，由逐次逼近法，总有：当 $y_0\in U(0,\delta)$ 时，$\exists x_1\in B\left(0,\frac{1}{3}\right)$, s. t. $\|y_0-T(x_1)\|_Y<\frac{\delta}{3}\Rightarrow$ 令 $y_1=y_0-T(x_1)\in U\left(0,\frac{\delta}{3}\right)$, $\exists x_2\in B\left(0,\frac{1}{3^2}\right)$, s. t. $\|y_1-T(x_2)\|_Y<\frac{\delta}{3^2}\Rightarrow\cdots\Rightarrow$ 令 $y_n=y_{n-1}-T(x_n)\in U\left(0,\frac{\delta}{3^n}\right)$, $\exists x_{n+1}\in B\left(0,\frac{1}{3^{n+1}}\right)$, s. t. $\|y_n-T(x_{n+1})\|_Y<\frac{\delta}{3^{n+1}}\Rightarrow$ 继续此步骤，最终得 $\sum_{n=1}^{+\infty}\|x_n\|_X\leqslant\sum_{n=1}^{+\infty}\frac{1}{3^n}=\frac{1}{2}$. 令 $x_0\equiv\sum_{n=1}^{+\infty}x_n$, $\|x_0\|_X\leqslant\frac{1}{2}$, 即 $x_0\in B(0,1)$. 记 $s_n=\sum_{j=1}^{n}x_j\Rightarrow$ 由前一步 $y_n=y_{n-1}-T(x_n)\in U\left(0,\frac{\delta}{3^n}\right)$ 蕴含 $\frac{\delta}{3^n}>\|y_n\|_Y=\|y_{n-1}-T(x_n)\|_Y=\cdots=\|y_0-T(x_1+x_2+\cdots+x_n)\|_Y=\left\|y_0-T\left(\sum_{j=1}^{n}x_j\right)\right\|_Y\Rightarrow\left\|y_0-T\left(\sum_{j=1}^{n}x_j\right)\right\|_Y<\frac{\delta}{3^n}\Rightarrow\|y_0-T(s_n)\|_Y<\frac{\delta}{3^n}$, $n=1,2,\cdots\Rightarrow\lim\limits_{n\to+\infty}T(s_n)=y_0$.

为完成第二步证明，利用范数的连续性与 T 的连续性证 $y_0=T(x_0)$.

由 $\|y_0-T(x_0)\|_Y=\|y_0-T(s_n)+T(s_n)-T(x_0)\|_Y\leqslant\|y_0-T(s_n)\|_Y+\|T(s_n)-T(x_0)\|_Y=\|y_0-T(s_n)\|_Y+\|T(s_n-x_0)\|_Y\Rightarrow$ 对于任给 $\varepsilon>0$，利用 $\|T\|\leqslant M$，由 $x_0=\sum_{n=1}^{+\infty}x_n$ 蕴含 $\forall\varepsilon>0$, $\exists N_1>0$, s. t. $n>N_1$ 时，成立 $\|s_n-x_0\|_X<\frac{\varepsilon}{2M}$；且 $\exists N_2>0$, s. t. 当 $n>N_2$ 时，成立 $\|T(s_n)-T(x_0)\|_Y<M\|s_n-x_0\|_X<\frac{\varepsilon}{2}$；进而，$\exists N_3>0$, s. t. 当 $n>N_3$ 时，$\frac{\delta}{3^n}<\frac{\varepsilon}{2}\Rightarrow\forall\varepsilon>0,\exists N>0$, s. t. $n>N$ 时，$\|y_0-T(s_n)\|_Y<\frac{\varepsilon}{2}\Rightarrow\|y_0-T(x_0)\|_Y\leqslant\|y_0-T(s_n)\|_Y+\|T(s_n-x_0)\|_Y<\varepsilon\Rightarrow y_0=T(x_0)$.

于是，定理得证.

2）逆算子定理

定理 4.2.3 设 $T:X\to Y$ 是 Banach 空间 X 到 Y 的有界线性算子，且 T 是一一满射，则 T 存在逆映射 T^{-1}，并且 $T^{-1}:Y\to X$ 是 Y 到 X 的有界线性算子；亦即，

$$T\in\mathfrak{B}(X,Y)\xrightarrow{T(X)=Y,\text{一对一}}T^{-1}\in\mathfrak{B}(Y,X).$$

证 由假设，$T:X\to Y$ 是 X 到 Y 上的有界线性算子，$T\in\mathfrak{B}(X,Y)$，且 T 是一一的满射，

故 $T^{-1}:Y\to X$ 存在，也是一一的满射. 于是，只要证明逆映射 T^{-1} 是线性的，并且 $T^{-1}\in\mathfrak{B}(Y,X)$.

(1) 证明 T^{-1} 的线性.

$$\begin{aligned}
& y_1,y_2\in Y\Rightarrow \exists x_1,x_2\in X,\ \text{s.t. } y_1=T(x_1),y_2=T(x_2)\\
\Rightarrow\ & x_1=T^{-1}(y_1),x_2=T^{-1}(y_1)\\
\Rightarrow\ & T^{-1}(y_1)+T^{-1}(y_2)=x_1+x_2\\
\Rightarrow\ & T(T^{-1}(y_1)+T^{-1}(y_2))=T(x_1+x_2)=T(x_1)+T(x_2)=y_1+y_2\\
\Rightarrow\ & T^{-1}(y_1)+T^{-1}(y_2)=T^{-1}(y_1+y_2);\\
& y\in Y,\alpha\in\mathbb{F}\Rightarrow\exists x\in X,\ \text{s.t. } y=T(x)\\
\Rightarrow\ & T^{-1}(y)=x\Rightarrow T(\alpha T^{-1}(y))=T(\alpha x)=\alpha T(x)=\alpha y\\
\Rightarrow\ & \alpha T^{-1}(y)=T^{-1}(\alpha y).
\end{aligned}$$

(2) 证明 $T^{-1}\in\mathfrak{B}Y,X)$.

利用开映射定理中的 $T(B(0,1))\supset U(0,\delta)$，并改写为 $U(0,1)\subset T\left(B\left(0,\frac{1}{\delta}\right)\right)\Rightarrow$ $T^{-1}(U(0,1))\subset B\left(0,\frac{1}{\delta}\right)\Rightarrow\forall y\in Y,\ \|y\|_Y<1$ 有 $\|T^{-1}(y)\|_Y<\frac{1}{\delta}\Rightarrow$ 由算子的线性，$\forall y\in X,\forall\varepsilon>0$ 蕴含 $\|T^{-1}(y)\|_Y<\frac{(1+\varepsilon)}{\delta}\|y\|_Y\Rightarrow$ 令 $\varepsilon\to0$，得 $\forall y\in Y$，有 $\|T^{-1}(y)\|_X<\frac{1}{\delta}\|y\|_Y\Rightarrow T^{-1}\in\mathfrak{B}Y,X)$.

3) 闭图像定理

定义 4.2.5(闭算子) 设 $T:\mathfrak{D}\to Y$ 是 X 的子集 $\mathfrak{D}\subset X$ 到 Y 的线性算子，称

$$G(T)=\{(x,T(x))\in X\times Y:x\in\mathfrak{D}\}$$

为**算子 T 在 $X\times Y$ 中的图(graph)**. 若 $G(T)$ 是 $X\times Y$ 中的闭集，则称 T 是**闭线性算子**，(**closed linear operator**)，简称闭算子.

闭算子的定义等价于下面命题.

命题 4.2.1 设 $T:\mathfrak{D}\to Y$ 是线性算子. T 为闭算子，当且仅当 $\forall\{x_n\}\subset\mathfrak{D}$，若 $x_n\xrightarrow{X}x\in X$ 且 $Tx_n\xrightarrow{Y}y\in Y$，则有 $x\in\mathfrak{D}$ 与 $T(x)=y$；亦即 $(x,T(x))\in G(T)=\mathfrak{D}\times Y$.

对这一等价性的证明可以加深对闭算子的理解.

证 (1) 设 $T:\mathfrak{D}\to Y$ 是闭算子，亦即 $G(T)=\mathfrak{D}\times Y$ 是 $X\times Y$ 中的闭集 $\Rightarrow\forall\{x_n\}\subset\mathfrak{D}$，$x_n\xrightarrow{X}x\in X,Tx_n\xrightarrow{X}Tx\in Y\Rightarrow$ 在乘积拓扑下，$(x_n,T(x_n))\in\mathfrak{D}\times Y\xrightarrow{X\times Y}(x,T(x))\in X\times Y$ $\Rightarrow$ 因 $G(T)=\mathfrak{D}\times Y$ 是 $X\times Y$ 中的闭集，故 $(x,T(x))\in\mathfrak{D}\times Y$.

(2) 设 $(x,y)\in\overline{G(T)}=\overline{\mathfrak{D}\times Y}\Rightarrow\exists\{x_n\}\subset\mathfrak{D},(x_n,T(x_n))\in G(T)\xrightarrow{X\times Y}(x,y)\in X\times$

Y,亦即 $\lim_{n\to+\infty}\|(x_n-x,T(x_n)-y)\|_{X\times Y}=0 \Rightarrow \|x_n-x\|_X\leqslant\|(x_n-x,T(x_n)-y)\|_{X\times Y}\to 0$,
$\|T(x_n)-y\|_Y\leqslant\|(x_n-x,T(x_n)-y)\|_{X\times Y}\to 0$,亦即 $\{x_n\}\subset\mathfrak{D}$, $x_n\xrightarrow{X}x\in X$ 且 $T(x_n)\xrightarrow{Y}y\in Y\Rightarrow$ 由充分性假设,这蕴含 $(x,T(x))\in\mathfrak{D}\times Y=G(T)\Rightarrow G(T)$是 $X\times Y$ 中的闭集.

定理 4.2.4 设 $T:\mathfrak{D}\to Y$ 是闭线性算子,且 $\mathfrak{D}\subseteq X$ 是闭集,则 T 是有界线性算子;亦即

$$T\text{ 是闭线性算子}\xrightarrow{\mathfrak{D}\text{ 是闭集}}T\in\mathfrak{B}(X,Y).$$

为了证明这个定理,先要引入范数强弱性概念与一个重要命题.

范数的强弱性(范数的比较) 若在同一空间 X 上有两个范数 $\|x\|_1$, $\|x\|_2$,使得 $(X,\|x\|_1)$,$(X,\|x\|_2)$都是赋范线性空间.称范数 $\|x\|_1$ 比 $\|x\|_2$强,若存在常数$M>0$,使得 $\|x\|_2\leqslant M\|x\|_1$;称 $\|x\|_1$ 与 $\|x\|_2$等价,若存在常数 $M>0$,$m>0$,使得 $m\|x\|_1\leqslant\|x\|_2\leqslant M\|x\|_1$.

命题 4.2.2 设$(X,\|x\|_1)$,$(X,\|x\|_2)$都是 Banach 空间,且范数 $\|x\|_1$ 比 $\|x\|_2$强,则 $\|x\|_1$ 与 $\|x\|_2$彼此等价.

证 考虑恒同映射 $I:X\to X$,它是$(X,\|x\|_1)$到$(X,\|x\|_2)$的线性算子,由于 $\|x\|_1$比 $\|x\|_2$强,故存在 $M>0$,使得

$$\|I(x)\|_2\leqslant M\|x\|_1, \tag{4.2.6}$$

由此式知,$I:(X,\|x\|_2)\to(X,\|x\|_1)$是连续映射. 显然 I 是单射、满射,由逆算子定理知,I^{-1}是单射、满射、有界线性映射,所以存在 $c>0$,使得 $\|I^{-1}(x)\|_1\leqslant c\|x\|_2$. 但是,$I(x)=x$, $I^{-1}(x)=x$,若令 $m=\dfrac{1}{c}$,则上式成为 $m\|x\|_1\leqslant\|x\|_2$,联合(4.2.6)式得 $m\|x\|_1\leqslant\|x\|_2\leqslant M\|x\|_1$,故 $\|x\|_1$ 与 $\|x\|_2$ 彼此等价.

下面来完成定理 4.2.4 的证明.

证 因 $X,Y,X\times Y$ 都是 Banach 空间,若 T 为闭算子,则 $G(T)=\mathfrak{D}\times Y$ 是 $X\times Y$ 中的闭集,也是闭线性子空间,从而 $G(T)=\mathfrak{D}\times Y$ 也是 Banach 空间,并且$(\mathfrak{D},\|x\|_X)$也是一个 Banach 空间.

现在,在 $\mathfrak{D}$ 上定义一个新范数 $\|x\|_D$,满足

$$\|x\|_D=\|x\|_X+\|T(x)\|_Y,\quad\forall x\in\mathfrak{D}, \tag{4.2.7}$$

我们证明,$(\mathfrak{D},\|x\|_D)$也是一个 Bananch 空间.

取基本列$\{x_n\}\subset(\mathfrak{D},\|x\|_D)$,由(4.2.7)式,有

$$\|x_n-x_m\|_D=\|x_n-x_m\|_X+\|T(x_n)-T(x_m)\|_Y\to 0,\ n,m\to+\infty;$$

因 X,Y 为 Banach 空间,故 $\exists x\in X$, $\exists y\in Y$,使得 $x_n\xrightarrow{X}x$, $T(x_n)\xrightarrow{Y}y\,(n\to+\infty)$;又由

T 为闭算子，得到 $x\in\mathfrak{D}$，$T(x)=y$. 于是，$T(x_n)\xrightarrow{Y}T(x)(n\to+\infty)$. 这蕴含

$$\|x_n-x\|_D=\|x_n-x\|_X+\|T(x_n)-T(x)\|_Y\to 0,\ n,m\to+\infty,$$

所以$(\mathfrak{D},\|x\|_D)$是 Banach 空间. 由(4.2.7)式有 $\|x\|_D=\|x\|_X+\|T(x)\|_Y\geqslant\|T(x)\|_Y$.

另一方面，也有 $\|x\|_D\geqslant\|x\|_X$，于是，利用范数强弱性概念及命题 4.2.2，知 $\exists M>0$，使得 $\|x\|_D\leqslant M\|x\|_X$.

这样，对于 $x\in X_1$，有 $\|T(x)\|_Y\leqslant\|x\|_D\leqslant M\|x\|_X$，所以，$T:X\to Y$ 是 X 到 Y 上的有界线性算子. 闭图像定理得证.

3. 算子空间 $\mathfrak{B}(X,Y)$中的一些重要性质

1）有界线性算子空间$(\mathfrak{B}(X,Y),\|T\|)$中的收敛性

（1）按算子范数收敛性（强收敛性）

定义 4.2.6（算子序列按算子范数的收敛性） 设 $T,T_n\in\mathfrak{B}(X,Y)$，若

$$\lim_{n\to+\infty}\|T_n-T\|=0,$$

则称**算子序列$\{T_n\}$按算子范数收敛于算子** T；亦即，赋范线性空间$(\mathfrak{B}(X,Y),\|T\|)$中的元的序列$\{T_n\}$在算子范数$\|T\|$意义下的收敛性，也称算子序列$\{T_n\}$**强收敛于算子** T.

算子序列按算子范数收敛的等价定理叙述如下.

定理 4.2.5 算子序列$\{T_n\}$按算子范数收敛于算子 T，等价于$\{T_n\}$在 X 中的"单位球"

$$S=\{x\in X:\|x\|_X=1\}$$

上一致收敛.

证 由 $\lim\limits_{n\to+\infty}\|T_n-T\|=0\Leftrightarrow\lim\limits_{n\to+\infty}\sup\limits_{\|x\|_X=1}\|T_n(x)-T(x)\|_Y=0$，这也等价于

$$\lim_{n\to+\infty}\|T_n(x)-T(x)\|_Y=0,\quad\forall x\in S\text{ 一致成立}.$$

实际上，也等价于$\{T_n\}$在 X 中的任意有界闭集上的一致收敛性.

（2）按点收敛性

定义 4.2.7（算子序列按点的收敛性） 设 $T,T_n\in\mathfrak{B}(X,Y)$，若

$$\lim_{n\to+\infty}T_n(x)=T(x),\quad\forall x\in X,$$

则称**算子序列$\{T_n\}$按点收敛于算子** T，这里的极限是按照 Y 的范数成立的，亦即

$$\lim_{n\to+\infty}\|T_n(x)-T(x)\|_Y=0,\quad\forall x\in X.$$

2）有界线性算子空间 $\mathfrak{B}(X,Y)$的性质

定理 4.2.6（$\mathfrak{B}(X,Y)$的完备性定理） 设$(X,\|x\|_X)$为赋范线性空间，$(Y,\|y\|_Y)$为 Banach 空间，则算子空间$(\mathfrak{B}(X,Y),\|T\|)$是 Banach 空间，这里的$\|T\|$是算子范数.

注 当$(X,\|x\|_X)$，$(Y,\|y\|_Y)$为赋范线性空间时，$\mathfrak{B}(X,Y)$未必是 Banach 空间，但只要

$(Y, \|y\|_Y)$是 Banach 空间，$\mathfrak{B}(X,Y)$就必定是 Banach 空间，不论$(X, \|x\|_X)$是否是 Banach 空间.

定理 4.2.6 的证明留给读者.

定理 4.2.7($\mathfrak{B}(X,Y)$中的共鸣定理) 设$(X, \|x\|_X)$，$(Y, \|y\|_Y)$是 Banach 空间，$W=\{T\}\subset\mathfrak{B}(X,Y)$是有界线性算子族，若对于每个 $x\in X$，有$\sup\limits_{T\in W}\{\|T(x)\|_Y\}<+\infty$，则$V=\{\|T\|\}$是$\mathbb{R}$中的有界集. 亦即，$\exists M>0$，使得$\|T\|\leqslant M, \forall T\in W$.

分析 (1) 共鸣定理的条件"对于有界线性算子族 $W=\{T\}\subset\mathfrak{B}(X,Y)$，若每个 $x\in X$ 都有$\sup\limits_{T\in W}\{\|T(x)\|_Y\}<+\infty$"可理解为"对于有界线性算子族 $W=\{T\}$，$\forall x\in X$，$\exists M_x>0$，s. t. $\forall T\in W$，有$\|T(x)\|_Y\leqslant M_x$，此 M_x 与 T 无关"，亦即，"算子族 $W\subset\mathfrak{B}(X,Y)$中的每个算子 $T\in W$，在每个 $x\in X$ 上，有公共的(有限)上确界 $\sup\limits_{T\in W}\{\|T(x)\|_Y\}$".

定理结论为"$\exists M>0$，s. t. $\forall T\in W$ 有 $\|T\|\leqslant M$". 亦即，$\{\|T\|:T\in W\}\subset\mathbb{R}$ 视为$\mathbb{R}$中的子集是有界的，或$\{\|T\|:T\in W\}\subset\mathbb{R}$ 关于$\forall T\in W$ 是一致有界的. 因此，共鸣定理是说：**"算子族** $W=\{T\}\subset\mathfrak{B}(X,Y)$对于 $x\in X$ **点点有界"蕴含"算子族** W **对** $T\in W$ **一致有界"**，故共鸣定理又称为**一致有界原理**.

(2) 如果从逆否命题的角度来叙述本定理，则是"$W=\{T\}\subset\mathfrak{B}(X,Y)$，$\sup\limits_{T\in W}\|T\|=+\infty\Rightarrow\exists x_0\in X$，s. t. $\sup\limits_{T\in W}\|T(x_0)\|_Y=+\infty$"，故此定理被称为"**共鸣定理**".

(3) 由"$\forall x\in X$，$\sup\limits_{T\in W}\{\|T(x)\|_Y\}<+\infty$"证明"$\exists M>0$，s. t. $\forall T\in W$ 有 $\|T(x)\|_Y\leqslant M\|x\|_X$" 的思路如下：

① 由于"$\forall x\in X$，$\sup\limits_{T\in W}\{\|T(x)\|_Y\}<+\infty$" $\Leftrightarrow$ "$\forall x\in X$，$\exists M_x>0$，s. t. $\forall T\in W$ 有$\|T(x)\|_Y\leqslant M_x\|x\|_X$"，故需要找$\{M_x:x\in X\}$的上确界，它是一个共同的、与 T 无关的数 $M>0$，使得

$$\|T(x)\|_Y\leqslant M\|x\|_X,\quad \forall T\in W. \tag{4.2.8}$$

但从$\|T(x)\|_Y\leqslant M_x\|x\|_X$求上确界却得不到什么. 因此解决问题的关键是如何把假设条件$\sup\limits_{T\in W}\{\|T(x)\|_Y\}<+\infty$用上去.

量 $\sup\limits_{T\in W}\{\|T(x)\|_Y\}$有两个特点，一是它只与 $x\in X$ 有关，对于每个 $T\in W$ 都是同一个值；二是$\forall x\in X$，都是一个非负有限值，可以用 $\sup\limits_{T\in W}\|T(x)\|_Y$来表征 $x\in X$：把它加到 $x\in X$ 的范数上得到$\|x\|_W=\|x\|_X+\sup\limits_{T\in W}\|T(x)\|_Y$. 若能证明$\|x\|_W$是 x 的一个"新范数"，且与$\|x\|_X$等价，即存在 $c_1,c_2>0$，使得不等式 $c_2\|x\|_X\leqslant\|x\|_W\leqslant c_1\|x\|_X$成立，则所需证明的不等式(4.2.8)就成为我们所需要的不等式

$$\|T(x)\|_Y\leqslant\sup_{T\in W}\|T(x)\|_Y\leqslant\|x\|_X+\sup_{T\in W}\|T(x)\|_Y=\|x\|_W\leqslant c_1\|x\|_X.$$

② 注意到，由"新范数"的定义，$\|x\|_W=\|x\|_X+\sup\limits_{T\in W}\|T(x)\|_Y\geqslant\|x\|_X$，知$\|x\|_W$比$\|x\|_X$强，利用 Banach 空间中的命题 4.2.1，只要证明$(X, \|x\|_W)$是一个 Banach 空间就够了.

现在可以证明共鸣定理了.

证 $\forall x\in(X,\|x\|_X)$,定义一个新范数 $\|x\|_W=\|x\|_X+\sup\limits_{T\in W}\|T(x)\|_Y$. 不难看出,$\|x\|_W$是 X 上的一个范数,且比 $\|x\|_X$强. 故只要证明$(X,\|x\|_W)$是完备的.

事实上,取关于范数 $\|x\|_W$的基本列$\{x_n\}\subset X$, $\|x_n-x_m\|_W\to 0(n,m\to+\infty)$. 因

$$\|x_n-x_m\|_W=\|x_n-x_m\|_X+\sup_{T\in W}\|T(x_n-x_m)\|_Y,$$

得到

$$\|x_n-x_m\|_X\to 0,\quad n,m\to+\infty$$

与

$$\sup_{T\in W}\|T(x_n-x_m)\|_Y\to 0,\quad n,m\to+\infty.$$

于是,由$(X,\|x\|_X)$的完备性,$\exists x\in X$, s. t. $x_n\xrightarrow{X}x\,(n\to+\infty)$,即 $\forall\varepsilon>0$, $\exists N>0$, s. t. 当 $n>N$ 时, $\|x_n-x\|_X<\varepsilon$,以及 $\sup\limits_{T\in W}\|T(x_n-x)\|_Y<\varepsilon$(此式可由在 $\sup\limits_{T\in W}\|T(x_n-x_m)\|_Y$中,令 $m\to+\infty$得到).

又由假设,$\forall x\in X$,成立 $\sup\limits_{T\in W}\{\|T(x)\|_Y\}<+\infty$,于是,对于基本列$\{x_n\}\subset X$ 及其极限 $x\in X$,得到 $\forall T\in W$,有 $\|x_n-x\|_W=\|x_n-x\|_X+\sup\limits_{T\in W}\|T(x_n-x)\|_Y\to 0(n\to+\infty)$,即空间$(X,\|x\|_W)$是完备的.

然后,利用命题 4.2.1,知范数 $\|x\|_W$与 $\|x\|_X$等价. 因此,存在常数 $M>0$,使得

$$\|T(x)\|_Y\leqslant\sup_{T\in W}\|T(x)\|_Y\leqslant\|x\|_W\leqslant M\|x\|_X,\quad\forall x\in X,$$

这蕴含 $\forall T\in W$,算子范数 $\|T\|\leqslant M$. 共鸣定理得证.

作为共鸣定理的应用,另一个有用的结果是 $\mathfrak{B}(X,Y)$中的 **Banach-Steinhaus 定理**.

定理 4.2.8(Banach-Steinhaus 定理) 设$(X,\|x\|_X)$与$(Y,\|y\|_Y)$是 Banach 空间,$A\subset X$ 是 X 的稠密子集,$X=\overline{A}$. 设 $T,T_n\in\mathfrak{B}(X,Y)$,则

$$\lim_{n\to+\infty}T_n(x)=T(x),\quad\forall x\in X,\tag{4.2.9}$$

当且仅当(1) $\|T_n\|$ 有界;(2) $\lim\limits_{n\to+\infty}T_n(x)=T(x)$, $\forall x\in A$.

证 必要性 由假设条件(4.2.9),知(2)成立. 为证(1),设 $W=\{T_n\}\subset\mathfrak{B}(X,Y)$,于是,由(4.2.9)式,知 $\lim\limits_{n\to+\infty}\|T_n(x)-T(x)\|_Y=0$, $\forall x\in X$. 故任给 $\varepsilon>0$,存在 $N>0$, s. t. $n>N$ 时,有

$$\|T_n(x)\|_Y<\|T(x)\|_Y+\varepsilon,\quad\forall x\in X,\quad n=1,2,\cdots,N.$$

由 $\|T_n(x)\|_Y<M_x\|x\|_X$, $\forall x\in X$, $n=1,2,\cdots,N$,故 $\forall x\in X$, $\sup\limits_{T_n\in W}\|T(x)\|_X<+\infty$. 据共鸣定理,得到 $\|T_n\|$ 有界,(1)得证.

充分性 假设(1)、(2)成立,我们证明(4.2.9)式.

由(1),设 $\|T_n\|\leqslant M$, $\forall n\in\mathbb{Z}^+$. 于是,$\forall x\in X$, $\forall\varepsilon>0$,由 $X=\overline{A}$,存在 $y\in A$,使得

$\|x-y\|_X \leqslant \frac{1}{4(\|T\|+M)}\varepsilon$；故 $\forall n>N$，有

$$\begin{aligned}\|T_n(x)-T(x)\|_Y &\leqslant \|T_n(x)-T_n(y)\|_Y+\|T_n(y)-T(y)\|_Y+\|T(y)-T(x)\|_Y \\ &\leqslant M\|x-y\|_X+\|T_n(y)-T(y)\|_Y+M\|y-x\|_X \\ &< \frac{\varepsilon}{4}+\|T_n(y)-T(y)\|_Y+\frac{\varepsilon}{4}=\frac{\varepsilon}{2}+\|T_n(y)-T(y)\|_Y.\end{aligned}$$

对 $\|T_n(y)-T(y)\|_Y$，由 $y\in A$ 与假设条件(2)，取足够大的 n，便得 $\|T_n(y)-T(y)\|_Y<\frac{\varepsilon}{2}$. 于是，$\forall x\in X, \forall \varepsilon>0, \exists N>0$, s. t. $n>N$ 时，$\|T_n(x)-T(x)\|_Y\leqslant\varepsilon$. 此即(4.2.9)式. 定理得证.

下面给出共鸣定理的两个重要应用.

例 4.2.4(求积分公式的收敛性) 对于熟悉的 R 积分的近似公式

$$\int_a^b f(x)\mathrm{d}x \approx \sum_{k=0}^{n} f(x_k)\Delta x_k,$$

所取分法为 $a=x_0<x_1<\cdots<x_{n-1}<x_n=b$. 需要考虑的问题是：当 $n\to+\infty$ 时，需要加什么条件，才能在使用上述近似公式所产生的误差趋向于零？

由于所取分点与 n 有关，于是 $\sum_{k=0}^{n} f(x_k)\Delta x_k \approx \sum_{k=0}^{n} A_k^{(n)} f(x_k^{(n)})$. 问题化为

$$\lim_{n\to+\infty}\sum_{k=0}^{n} A_k^{(n)} f(x_k^{(n)}) = \int_a^b f(x)\mathrm{d}x$$

在什么条件下成立？

命题 4.2.3(求积公式的收敛性条件) 对于每个连续函数 $f(x)\in C([a,b])$，公式

$$\lim_{n\to+\infty}\sum_{k=0}^{n} A_k^{(n)} f(x_k^{(n)}) = \int_a^b f(x)\mathrm{d}x \tag{4.2.10}$$

成立，当且仅当以下两个条件成立：

(1) 存在常数 $M>0$，使得 $\sum_{k=0}^{n} |A_k^{(n)}| \leqslant M$；

(2) $\lim_{n\to+\infty}\sum_{k=0}^{n} A_k^{(n)} f(x_k^{(n)}) = \int_a^b f(x)\mathrm{d}x$ 对于每个多项式

$$f(x)=p(x)=a_0+a_1x+\cdots+a_mx^m,\quad x\in[a,b],\ m\in\mathbb{Z}^+$$

成立.

证 ① 证明 $T_n\in\mathfrak{B}(C([a,b]),\mathbb{R})$. 考虑 Banach 空间 $C([a,b])$ 上的泛函

$$T_n(f)=\sum_{k=0}^{n} A_k^{(n)} f(x_k^{(n)}),\quad f(x)\in C([a,b]),\quad x\in[a,b],$$

则(4.2.10)式可写为

$$\lim_{n\to+\infty} T_n(f) = \int_a^b f(x)\mathrm{d}x = T(f),\quad f\in C([a,b]).$$

于是，$T_n(f):C([a,b])\to\mathbb{R}$.

$T\in L(C([a,b]),\mathbb{R})$显然，下证 $T_n\in\mathfrak{B}(C([a,b]),\mathbb{R})$.

由于$\forall f\in C([a,b])$，有 $\|f\|_{C([a,b])}\leqslant M_1$，故

$$|T_n(f)|=\left|\sum_{k=0}^{n}A_k^{(n)}f(x_k^{(n)})\right|\leqslant\left(\sum_{k=0}^{n}|A_k^{(n)}|\right)\|f\|_{C([a,b])}\leqslant M_1\left(\sum_{k=0}^{n}|A_k^{(n)}|\right),$$

因此，$\|T_n\|\leqslant\left(\sum_{k=0}^{n}|A_k^{(n)}|\right)\leqslant M_1M=M_2$.

另一方面，对于每个 $n\in\mathbb{N}$，取 $f_n(x)\in C([a,b])$，不失一般性，使得

$$f_n(x_k^{(n)})=\operatorname{sgn}A_k^{(n)},\quad k=0,1,\cdots,n,$$

且满足 $\|f_n\|_{C([a,b])}=\max\{|f_n(x)|:x\in[a,b]\}=1$.

于是，$\|T_n\|\geqslant|T_n(f_n)|=\sum_{k=0}^{n}|A_k^{(n)}|$，从而

$$\|T_n\|=\sum_{k=0}^{n}|A_k^{(n)}|,\quad n\in\mathbb{N}.\tag{4.2.11}$$

所以，$T_n\in\mathfrak{B}(C([a,b]),\mathbb{R})$. 并且，条件(1)就是 $\|T_n\|\leqslant M,\forall n\in\mathbb{N}$.

② 证明充分性. 设条件(1)、(2)成立，则由于多项式全体所成的集合在 $C([a,b])$中稠密，故极限

$$\lim_{n\to+\infty}\sum_{k=0}^{n}A_k^{(n)}f(x_k^{(n)})=\int_a^b f(x)\mathrm{d}x$$

对每个 $f\in C([a,b])$成立. 充分性得证.

③ 证明必要性. 设(4.2.10)式成立，即

$$\lim_{n\to+\infty}\sum_{k=0}^{n}A_k^{(n)}f(x_k^{(n)})=\lim_{n\to+\infty}T_n(f)=T(f)=\int_a^b f(x)\mathrm{d}x,\quad f\in C([a,b])$$

成立，亦即

$$\lim_{n\to+\infty}|T_n(f)-T(f)|=0,\quad\forall f\in C([a,b]),$$

故$\sup\limits_n|T_n(f)-T(f)|\leqslant M$(一致收敛数列必有界)，再由 Banach-Steinhaus 定理，知 $\|T_n\|\leqslant M$，所以(1)成立，当然由(4.2.11)式立即得到(2).

例 4.2.5(Fourier 级数的敛散性) 设 $C_{2\pi}$是以 2π 为周期的实值连续函数的全体所成的 Banach 空间，$f\in C_{2\pi}$的范数定义为 $\|f\|_{C_{2\pi}}=\max\limits_{x\in[0,2\pi]}|f(x)|$.

对于熟悉的 Fourier 级数

$$f(x)\sim\frac{a_0}{2}+\sum_{k=1}^{+\infty}(a_k\cos kx+b_k\sin kx),$$

作前 $n+1$ 项部分和

$$S_n(f)(x)=\frac{a_0}{2}+\sum_{k=1}^{n}(a_k\cos nx+b_k\sin nx)$$

$$= \frac{1}{\pi}\int_{\pi_0}^{2} f(x)\left(\frac{1}{2}+\sum_{k=1}^{n}\cos k(t-x)\right)\mathrm{d}t = \int_0^{2\pi} f(x)D_n(t,x)\mathrm{d}t,$$

其中 $D_n(t,x)=\dfrac{\sin\left(n+\frac{1}{2}\right)(t-x)}{\sin\frac{1}{2}(t-x)}$ $(n\in\mathbb{N})$ 称为 Dirichlet 核. 于是，$S_n(f): C_{2\pi}\to C_{2\pi}$，$\{S_n(f)\}_{n\in\mathbb{N}}$ 成为 $C_{2\pi}$ 到自身的线性算子族.

我们证明，$\forall x_0\in[0,2\pi]$，$\exists f\in C_{2\pi}$，使得 $f(x)$ 的 Fourier 级数在 x_0 发散.

事实上，可将 $[0,2\pi]$ 视为 $[-\pi,\pi]$，并且不失一般性，可设 $x_0=0$. 于是，作 $C_{2\pi}$ 上的线性泛函序列

$$T_n(f)=\int_{-\pi}^{\pi} f(t)D_n(t,0)\mathrm{d}t: C_{2\pi}\to\mathbb{R},$$

其中 $D_n(t,0)=\dfrac{1}{2\pi}+\dfrac{1}{\pi}\sum\limits_{k=1}^{n}\cos kt=\dfrac{\sin\left(n+\frac{1}{2}\right)t}{\sin\frac{1}{2}t}$ $(n\in\mathbb{N})$. 显然，$D_n(t,0)$ 是 $t\in[-\pi,\pi]$ 的连续函数，因此 $\|T_n\|\leqslant M$，并且可以证明，算子范数为 $\|T_n\|=\int_{-\pi}^{\pi}|D_n(t,0)|\mathrm{d}t, n\in\mathbb{N}$. 由于

$$\int_{-\pi}^{\pi}|D_n(t,0)|\mathrm{d}t=\int_0^{2\pi}|D_n(t,0)|\mathrm{d}t=\frac{1}{2\pi}\int_0^{2\pi}\frac{\left|\sin\left(n+\frac{1}{2}\right)t\right|}{\left|\sin\frac{1}{2}t\right|}\mathrm{d}t=\frac{1}{\pi}\int_0^{\pi}\frac{|\sin(2n+1)t|}{|\sin t|}\mathrm{d}t$$

$$=\frac{1}{\pi}\int_0^{\pi}\frac{|\sin(2n+1)t|}{\sin t}\mathrm{d}t\geqslant\frac{1}{\pi}\int_0^{\pi}\frac{|\sin(2n+1)t|}{t}\mathrm{d}t=\frac{1}{\pi}\int_0^{(2n+1)\pi}\frac{|\sin u|}{u}\mathrm{d}u,$$

最后一项当 $n\to+\infty$ 时是趋向于 $+\infty$ 的，因为反常积分 $\int_0^{+\infty}\frac{\sin t}{t}\mathrm{d}t$ 是非绝对收敛的. 于是，$\|T_n\|\to+\infty$.

这样，由 Banach-Steinhaus 定理知，至少存在一个元 $f_0\in C_{2\pi}$，使得

$$\{T_n(f_0)\}=\left\{T_n(f_0)=\int_{-\pi}^{\pi} f_0(t)D_n(t,0)\mathrm{d}t\right\}$$

发散，即 $\lim\limits_{n\to+\infty}T_n(f_0)=\infty$. 这是因为，若 $\forall f\in C_{2\pi}\Rightarrow\lim\limits_{n\to+\infty}T_n(f_0)=A\Rightarrow\|T_n(f)\|\leqslant M$，$\forall f\in C_{2\pi}$ 一致成立，从而导致 $\|T_n\|$ 一致有界，这与 $\|T_n\|\to+\infty$ 矛盾.

这个结果澄清了人们最初认为函数 $f\in C_{2\pi}$ 的 Fourier 级数一定收敛的错觉，从而促使人们研究级数的敛散性，发展了 Fourier 分析理论.

3) 有界线性算子空间 $\mathfrak{B}(X,X)\equiv\mathfrak{B}(X)$ 上的算子代数

在 $(X,\|x\|_X)$ 是赋范线性空间、$(Y,\|y\|_Y)$ 是 Banach 空间的假设下，在有界线性算子集合 $\mathfrak{B}(X,Y)$ 中，赋予加法运算 $+$、数乘运算 $\alpha\cdot$、范数 $\|T\|$，则 $(\mathfrak{B}(X,Y),+,\alpha\cdot,\|T\|)$ 成为 Banach 空间. 当 $\mathfrak{B}(X)\equiv\mathfrak{B}(X,X)$ 时，还可以考虑任意两个线性算子：$T,S\in\mathfrak{B}(X)$ 的复合 $S\circ T$，定义为 $(S\circ T)(x)=S(T(x))$，$\forall x\in X$，使得算子空间具有一种新的结构——算子代数（$\mathfrak{B}$

$(X,X),+,\alpha\cdot,\|T\|,S\circ T)$.

我们对于一般的赋范线性空间$(X,+,\alpha\cdot,\|x\|_X)$定义代数,然后再回到算子空间 $\mathscr{B}(X)$上来.(请读者比较定义 2.2.6.)

定义 4.2.8(**代数、赋范代数、具有单位元的赋范代数、Banach 代数**)

(1) **代数** 设$(X,+,\alpha\cdot)$是数域 $\mathbb{F}$ 上的线性空间,若对于任意元 $x,y,z,\cdots\in X$, $\alpha\in\mathbb{F}$,定义元素之间的"乘法"运算 $x\circ y$,满足

① $x\circ y\in X$, (封闭性)

② $(x\circ y)\circ z=x\circ(y\circ z)$, (结合律)

③ $x\circ(y+z)=x\circ y+x\circ z$, $(y+z)\circ x=y\circ x+z\circ x$, (加乘分配律)

④ $\alpha(x\circ y)=(\alpha x)\circ y=x\circ(\alpha y)$, (数乘结合律)

则称$(X,+,\alpha\cdot)$是一个**代数**,并记为$(X,+,\alpha\cdot,x\circ y)$.

进而,若乘法运算满足交换律

⑤ $x\circ y=y\circ x$, (交换律)

则称$(X,+,\alpha\cdot,x\circ y)$是一个**交换代数**.

(2) **赋范代数** 设$(X,+,\alpha\cdot,\|x\|_X)$是数域$\mathbb{F}$ 上的赋范线性空间,若在其中定义关于元之间的乘法 $x\circ y$,满足上述①～④,且范数满足

⑥ $\|x\circ y\|_X\leqslant\|x\|_X\|y\|_X$, (范数乘法不等式)

则称它是一个**赋范代数**,记为$(X,+,\alpha\cdot,\|x\|_X,x\circ y)$;若乘法可交换,则称其为**交换赋范代数**(亦即满足上述①～⑥).

(3) **具有单位元的赋范代数** 设$(X,+,\alpha\cdot,\|x\|_X,x\circ y)$是一个赋范代数,若存在一个元 $e\in X$,满足

⑦ $e\circ x=x\circ e=x$, $\forall x\in X$;且 $\|e\|_X=1$, (存在单位元)

则称 $e\in X$ 为 X 的单位元,并称$(X,+,\alpha\cdot,\|x\|_X,x\circ y,e)$是**具有单位元的赋范代数**;若乘法可交换,则称其为**具有单位元的交换赋范代数**.

(4) **Banach 代数** 设$(X,+,\alpha\cdot,\|x\|_X, x\circ y)$是一个赋范代数,且关于范数 $\|x\|_X$是一个 Banach 空间,则称它为一个 **Banach 代数**.

相应地有交换 Banach 代数、具有单位元的 Banach 代数、具有单位元的交换 Banach 代数等.

例 4.2.6 $C([a,b])$是一个具有单位元的交换 Banach 代数.

事实上,$C([a,b])$中的运算如下:

加法 $f,g\in C([a,b])\Rightarrow(f+g)(x)=f(x)+g(x)$, $\forall x\in[a,b]$;

数乘 $f\in C([a,b])$, $\alpha\in\mathbb{F}\Rightarrow(\alpha f)(x)=\alpha f(x)$, $\forall x\in[a,b]$;

乘法 $f,g\in C([a,b])\Rightarrow(f\cdot g)(x)=f(x)g(x)$, $\forall x\in[a,b]$;

乘法单位元 $f\equiv 1\in C([a,b])\Rightarrow(1\cdot f)(x)=1\cdot f(x)=f(x),\ \forall x\in[a,b]$;

范数 $f\in C([a,b])\Rightarrow\|f\|_{C([a,b])}=\max\limits_{x\in[a,b]}|f(x)|$;

因此,$(C([a,b]),+,\alpha\cdot,f\cdot g,\|f\|_{C([a,b])},1)$是一个**具有单位元的交换 Banach 代数**.

例 4.2.7 $W([0,2\pi])=\left\{f:f(x)=\sum\limits_{k=-\infty}^{+\infty}\xi_k e^{ikx},\sum\limits_{k=-\infty}^{+\infty}|\xi_k|<+\infty,x\in[0,2\pi]\right\}$ 是一个交换 Banach 代数,其中 $i=\sqrt{-1}$.

事实上,$W([0,2\pi])$的加法、数乘、范数都与例 4.2.6 相似,而“乘法”是如下定义的“卷积”.

设 $f,g\in W([0,2\pi])$,对于 $f(x)=\sum\limits_{k=-\infty}^{+\infty}\xi_k e^{ikx}$,$g(x)=\sum\limits_{k=-\infty}^{+\infty}\eta_k e^{ikx}$,乘法定义为卷积

$$(f*g)(x)=\sum_{n=-\infty}^{+\infty}\sum_{k=-\infty}^{+\infty}\xi_{n-k}\eta_k e^{inx}.$$

由估计

$$\begin{aligned}\|(f*g)(x)\|_{W([0,2\pi])}&=\sum_{n=-\infty}^{+\infty}\left|\sum_{k=-\infty}^{+\infty}\xi_{n-k}\eta_k\right|\leqslant\sum_{n=-\infty}^{+\infty}\sum_{k=-\infty}^{+\infty}|\xi_{n-k}||\eta_k|\\&=\sum_{k=-\infty}^{+\infty}\left(\sum_{n=-\infty}^{+\infty}|\xi_{n-k}|\right)|\eta_k|=\|f\|_{W([0,2\pi])}\|g\|_{W([0,2\pi])},\end{aligned}$$

知 $W([0,2\pi])$成为一个交换 Banach 代数.

例 4.2.8 设$(X,\|x\|_X)$是 Banach 空间,考虑有界线性算子空间 $\mathfrak{B}(X)\equiv\mathfrak{B}(X,X)$,它是一个 Banach 空间. 算子 $T,S\in\mathfrak{B}(X)$的乘法 $S\circ T$ 定义为算子复合$(S\circ T)(x)=S(T(x))$,$\forall x\in X$;恒同算子 I 是乘法的单位元. 于是,$(\mathfrak{B}(X),+,\alpha\cdot,\|T\|,S\circ T,I)$ 是一个**具有单位元的、不可交换的 Banach 代数**.

例 4.2.9 Banach 代数 $\mathfrak{A}(D)$ 设 $D=\{z\in\mathbb{C}:|z|\leqslant 1\}$,$\mathfrak{A}(D)$是 D 上定义的、D 内解析的复变量复值函数的全体,其中各种运算与例 4.2.6 类同,使得 $\mathfrak{A}(D)$成为一个**具有单位元的交换 Banach 代数**.

例 4.2.10 Banach 代数 $L([a,b])$ $[a,b]$区间上的 Lebesgue 可积函数空间 $L([a,b])$,加法、数乘与例 4.2.6 中的运算相同,范数定义为 $\|f\|_{L([a,b])}=\int_{[a,b]}|f(x)|dx$,乘法定义为卷积

$$f,g\in L([a,b])\Rightarrow(f*g)(x)=\int_{[a,b]}f(x-t)g(t)dt,$$

于是,$(L([a,b]),+,\alpha\cdot,\|f\|_{L([a,b])},f*g)$构成一个**没有单位元的、交换 Banach 代数**.

注 1 对于 $L([a,b])$中的卷积$(f*g)(x)=\int_{[a,b]}f(x-t)g(t)dt$,约定 $f,g\in L([a,b])$,当 $x\notin[a,b]$ 时,$f(x)=g(x)=0$. 卷积运算的封闭性由下述不等式得到:

$$f,g \in L([a,b]) \Rightarrow \|f * g\|_{L([a,b])} \leqslant \|f\|_{L([a,b])} \|g\|_{L([a,b])}.$$

注 2 Banach 代数 $L([a,b]) \equiv (L([a,b]),+,\alpha\cdot,\|f\|_{L([a,b])},f*g)$没有单位元的证明需要用到 Fourier 变换的知识(第 5 章,习题 11).

4.2.2 有界线性算子的谱理论

1. 线性算子的逆算子

逆算子(逆映射)的概念是大家熟悉的. 在算子谱理论的研究中,逆算子的存在与否成为研究的主题之一,并且有着广泛的应用. 首先来叙述一下逆算子的定义.

设$(X,\|x\|_X)$,$(Y,\|y\|_Y)$为 Banach 空间,$\mathfrak{B}(X,Y)$是 X 到 Y 的有界线性算子空间,对于 $T\in\mathfrak{B}(X,Y)$,若存在算子 $S:Y\to X$,使得算子的复合 $S\circ T:X\to X$ 是单位算子,即 $S\circ T=I:X\to X$,则称算子 $S:Y\to X$ 为算子 $T:X\to Y$ 的**逆算子(inverse operator)**,记为 $S=T^{-1}$,并称算子 T 为**可逆的(invertible)**.

注 1 在逆算子的定义中,实际上蕴含了算子 $T:X\to Y$ 是一一对应的.

注 2 若 $T_1:X\to Y$,$T_2:Y\to Z$ 是可逆算子,则 $T\equiv T_2\circ T_1:X\to Z$ 是可逆算子,且对于$(T_2\circ T_1)^{-1}:Z\to X$ 成立 $(T_2\circ T_1)^{-1}=T_1^{-1}\circ T_2^{-1}$.

注 3 对于 $\mathfrak{B}(X)\equiv\mathfrak{B}(X,X)$,两个可逆算子 $T_1:X\to X$、$T_2:X\to X$ 的复合是可逆的,且 $T_2\circ T_1:X\to X$ 满足$(T_2\circ T_1)^{-1}=T_1^{-1}\circ T_2^{-1}$.

2. 线性算子的谱理论

下面从大家熟悉的实数域$\mathbb{R}$ 上的有限维线性空间$(X,+,\alpha\cdot)$($\dim X=n$)出发,引入算子的谱理论.

我们知道,$\dim X=n \Rightarrow X \xleftrightarrow{\text{一对一}} \mathbb{R}^n$,即 X 与$\mathbb{R}^n$ 同构;而且

$$\mathfrak{B}(\mathbb{R}^n) \xleftrightarrow{\text{同构}} \mathfrak{M}_n,$$

这里 $\mathfrak{M}_n$是方阵空间,方阵 $A=[a_{ij}]_{n\times n}\in\mathfrak{M}_n$,视为算子 $T\leftrightarrow A=[a_{ij}]_{n\times n}$. 在高等数学教程中,求它的特征值问题是读者所熟悉的:如果 $\lambda\in\mathbb{C}$ 使得线性方程组

$$Ax=\lambda x \quad (\text{等价地},(A-\lambda I)x=0)$$

有非零解,其中 $x=\begin{bmatrix}x_1\\ \vdots\\ x_n\end{bmatrix}\in\mathbb{R}^n$,$\lambda\in\mathbb{R}$,则这样的 λ 称为算子 $T\leftrightarrow A=[a_{ij}]_{n\times n}$的特征值;所有特征值的全体称为算子 T 的**谱集**,简称**谱**. 特征值也称为**谱值**,或称为**谱点**,物理学中称为**本征值**. 记谱集为 $\sigma(T)\subset\mathbb{C}$. 复数 $\mu\in\mathbb{C}\setminus\sigma(T)$称为算子 T 的正则值.

于是,当算子 $T\leftrightarrow A=[a_{ij}]_{n\times n}$由方阵表征时,它的**谱就是方阵的特征值的全体所成的**

集 $\sigma(T)\subset\mathbb{C}$：$\sigma(T)=\{\lambda\in\mathbb{C}:(A-\lambda I)x=0\}$，**谱值是使得齐次线性方程组有非零解的** λ **值；正则值** $\mu\in\mathbb{C}\setminus\sigma(T)$ **是使得非齐次线性方程组有惟一解的** μ **值.**

由此可见，寻找一个算子的谱，与求方程的解的问题有密切联系. 把这个思路推广到有界线性算子上，就形成泛函分析的一个重要内容，**泛函方程(functional equation)**，特别是泛函微分方程的研究，更成为一个有意义、有实用价值的前沿课题.

定义 4.2.9(算子的正则值、预解式、谱值、谱) 设$(X,+,\alpha\cdot,\|x\|_X)$是数域$\mathbb{F}$上的 Banach 空间，若对于 X 上的任意有界线性算子 T，$T\in\mathfrak{B}(X)\equiv\mathfrak{B}(X,X)$，满足

(1) 复数 $\lambda\in\mathbb{C}$ 使得算子 $\lambda I-T$ 的逆算子$(\lambda I-T)^{-1}$存在，且$(\lambda I-T)^{-1}$定义在全空间 X 上，$(\lambda I-T)^{-1}:X\to X$，则称 $\lambda\in\mathbb{C}$ 为算子 T 的**正则值(regular value)**，并称 $R_\lambda=(\lambda I-T)^{-1}$为算子 T 的**预解式(resolvent)**；T 的正则值的全体，称为 T 的**正则集**，记为 $\rho(T)$.

(2) 若 $\lambda\in\mathbb{C}$ 不是算子 T 的正则值，亦即算子 $\lambda I-T$ 没有逆算子，则称 λ 为算子 T 的**谱值(spectrum value)**(或谱点)；算子 T 的谱值的全体，称为 T 的**谱集**，简称**谱(spectrum)**，记谱集为 $\sigma(T)$. $\sigma(T)\subset\mathbb{C}$ 中的谱点又分为两类：

① 对于 $\lambda\in\sigma(T)$，若齐次方程$(\lambda I-T)x=0$ 有非零解，则称 λ 为 T 的**特征值(eigen-value)**，或本征值，称对应于特征值 λ 的非零解 x 为**特征元(eigen-element)**或**特征向量(eigen-vector)**；记特征值全体所成的集为 $\sigma_p(T)$，我们有

$$\sigma_p(T)\subset\sigma(T)\subset\mathbb{C},$$

称 $\sigma_p(T)$为算子 T 的**点谱(point spectrum)**；

② 对于 $\lambda\in\sigma(T)$，若方程$(\lambda I-T)x=0$ 有且只有零解，亦即，对应于 λ 的特征向量只是零向量，而且算子 $\lambda I-T$ 的值域是 X 的真子空间，这样的 λ 的全体，记为 $\sigma_c(T)$，我们有

$$\sigma_c(T)\subset\sigma(T)\subset\mathbb{C},$$

称 $\sigma_c(T)$为算子 T 的**连续谱(continuous spectrum)**.

于是，有

$$\mathbb{C}=\rho(T)\cup\sigma(T)=\rho(T)\cup\{\sigma_p(T)\cup\sigma_c(T)\}.$$

例 4.2.11 在 n 维复空间$\mathbb{C}^n$ 中，三角阵

$$A=\begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ a_{n1} & 0 & \cdots & a_{nn} \end{bmatrix},\quad a_{ij}\in\mathbb{C},$$

定义一个算子

$$T:x=(\xi_1,\xi_2,\cdots,\xi_n)\in\mathbb{C}^n\to T(x)=(\eta_1,\eta_2,\cdots,\eta_n)\in\mathbb{C}^n,$$

使得 $\eta_k=\sum_{j=1}^n a_{kj}\xi_j\ (k=1,2,\cdots,n)$，亦即

$$T(x)=A[\xi_1\ \cdots\ \xi_n]^{\mathrm{T}}=\begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ a_{n1} & 0 & \cdots & a_{nn} \end{bmatrix}\begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix}=\begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_n \end{bmatrix}.$$

于是,算子 T 由三角阵 A 确定,而方程

$$0=(\lambda I-A)x=\begin{bmatrix} \lambda-a_{11} & 0 & \cdots & 0 \\ -a_{21} & \lambda-a_{22} & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ -a_{n1} & -a_{n2} & \cdots & \lambda-a_{nn} \end{bmatrix}x$$

的非零解正是对角线上的值 $\lambda=a_{11},a_{22},\cdots,a_{nn}$,也就是算子 T 的特征值,即

$$a_{11},a_{22},\cdots,a_{nn}\in\sigma_p(T),$$

其余的 $\lambda\in\mathbb{C}\setminus\sigma_p(T)=\rho(T)$ 便是 T 的正则值. 于是

$$\mathbb{C}=\rho(T)\cup\sigma_p(T).$$

例 4.2.12 对于 $[0,1]$ 上复值连续函数空间 $C([0,1])$,考虑乘法算子

$$(Tf)(x)=xf(x),\quad f\in C([0,1]),$$

知 $T:f\in C([0,1])\to T(f)(x)=xf(x)\in C([0,1])$ 是有界线性算子.

对于 $\lambda\notin[0,1]$,由

$$((\lambda I-T)(f))(x)=((\lambda-x)(f))(x),\quad f\in C([0,1]),$$

知 $\lambda I-T=\lambda-x$,预解式是 $R_\lambda(T)=(\lambda I-T)^{-1}=\dfrac{1}{\lambda-x}$,得

$$(R_\lambda f)(x)=((\lambda I-T)^{-1}f)(x)=\frac{1}{\lambda-x}f(x),$$

易于验证,R_λ 是定义在 $C([0,1])$ 上、取值为 $\dfrac{1}{\lambda-x}f(x)\in C([0,1])$ 的有界线性算子,且

$$(R_\lambda f)(x)\big|_{f(x)=(\lambda I-T)(f)}=((\lambda I-T)^{-1}(\lambda I-T)(f))(x)=f(x),$$

因此,$\lambda\notin[0,1]$ 是算子 T 的正则值.

对于 $\lambda\in[0,1]$,算子

$$((\lambda I-T)(f))(x)=(\lambda-x)f(x),\quad f\in C([0,1]).$$

当 $x=\lambda$ 时,$\lambda I-T=\lambda-x$ 无逆算子,但是 $(\lambda I-T)\big|_{x=\lambda}f(x)$ 为 0,因此不难证明,算子 $\lambda I-T$ 的值域是 $C([0,1])$ 的真子空间,并且齐次方程

$$(\lambda I-T)f(x)=0$$

在 $C([0,1])$ 中没有非零解,于是,$\lambda\in[0,1]$ 是算子 T 的连续谱中的点,$\lambda\in\sigma_c(T)=[0,1]$. 这样就有

$$\mathbb{C}=\rho(T)\cup\sigma_c(T).$$

例 4.2.13 对于复值连续函数空间 $C([0,1])$，考虑积分算子

$$(Tf)(x)=\int_0^x f(t)\mathrm{d}t,\quad f\in C([0,1]),$$

知 $T:f\in C([0,1])\to T(f)(x)=\int_0^x f(t)\mathrm{d}t\in C([0,1])$ 是有界线性算子.

对于 $\lambda\neq 0$，由

$$((\lambda I-T)(f))(x)=\lambda f(x)-\int_0^x f(t)\mathrm{d}t,\quad f\in C([0,1]),$$

知方程 $(\lambda I-T)f(x)=g(x)$ 就是方程

$$\lambda f(x)-\int_0^x f(t)\mathrm{d}t=g(x),$$

于是，当 $\lambda\neq 0$ 时，上式等价于

$$f(x)=\frac{1}{\lambda}g(x)+\frac{1}{\lambda}\int_0^x f(t)\mathrm{d}t.$$

由 Volterra 积分方程知识，此方程对于任何 $g(x)\in C([0,1])$ 存在惟一的解，这表明对于 $\lambda\neq 0$，有

$$(\lambda I-T)f(x)=g(x)\Leftrightarrow f(x)=(\lambda I-T)^{-1}g(x),$$

亦即算子 $(\lambda I-T)$ 可逆，所以，$\forall\lambda\neq 0\Rightarrow\lambda\in\rho(T)$.

当 $\lambda=0$ 时，齐次方程 $(\lambda I-T)f(x)=0$ 化为 $T(f(x))=-\int_0^x f(t)\mathrm{d}t=0$，由于 $f\in C([0,1])$，则 $T(f(x))=-\int_0^x f(t)\mathrm{d}t=0$ 蕴含 $f(t)=0$，因此相应的齐次方程只有零解，故 $\lambda=0$ 不是 T 的特征值.

另一方面，算子 $(\lambda I-T)\big|_{\lambda=0}$ 的值域中的元形如 $g(x)=-\int_0^x f(t)\mathrm{d}t$，满足 $g'(x)=-f(x)\in C([0,1])$，且 $g(0)=0$. 因此算子 T 的值域是 $W=\{g:g\in C^1([0,1]),g(0)=0\}$，而 W 是 $C([0,1])$ 的一个真子空间. 这表明 $\lambda=0\in\sigma_c(T)$，所以 $\sigma_c(T)=\{0\}$，$\sigma_p(T)=\varnothing$. 于是

$$\sigma(T)=\sigma_c(T),\quad \rho(T)=\mathbb{C}\setminus\{0\},\quad \mathbb{C}=\rho(T)\cup\sigma_c(T).$$

例 4.2.14 对于复值 L 可积函数空间 $L^1([0,1])$，考虑积分算子

$$(Tf)(x)=xf(x)+\int_x^1 f(t)\mathrm{d}t,\quad f\in L^1([0,1]),$$

知 $T:f\in L^1([0,1])\to T(f)(x)=xf(x)+\int_x^1 f(t)\mathrm{d}t\in L^1([0,1])$ 是有界线性算子.

对于 $\lambda\in(0,1]$，我们证明：区间 $[0,\lambda]$ 的示性函数 $\chi_{[0,\lambda]}(x)=\begin{cases}1, & x\in[0,\lambda],\\ 0, & x\notin[0,\lambda]\end{cases}$ 是对应于 λ 的特征向量. 事实上，对于 $0\leqslant x\leqslant\lambda$，有

$$(T\chi_{[0,\lambda]})(x)=x\chi_{[0,\lambda]}(x)+\int_x^1\chi_{[0,\lambda]}(t)\mathrm{d}t=x+\int_x^\lambda\mathrm{d}t=\lambda,$$

亦即，$((\lambda I-T)\chi_{[0,\lambda]})(x)=0$. 因此，$(0,1]$是算子 T 的点谱 $\sigma_p(T)=(0,1]$.

另一方面，当 $\lambda\notin[0,1]$时，任取 $g\in L^1([0,1])$，非齐次方程$((\lambda I-T)f)(x)=g(x)$化为

$$(\lambda-x)f(x)-\int_x^1 f(t)\mathrm{d}t=g(x),$$

令 $h(x)=\int_1^x f(t)\mathrm{d}t$，则得到一个具有初始条件的微分方程

$$\begin{cases}(\lambda-x)h'(x)+h(x)=g(x),\\ h(1)=0,\end{cases}$$

并且有惟一解

$$(\lambda-x)h'(x)+h(x)=(\lambda-x)f(x)+\int_1^x f(t)\mathrm{d}t=(\lambda-x)f(x)-\int_x^1 f(t)\mathrm{d}t=g(x).$$

这说明 $\lambda\notin[0,1]$时，算子$(\lambda I-T)$可逆，所以，$\forall\lambda\notin[0,1]\Rightarrow\lambda\in\rho(T)$.

当 $\lambda=0$ 时，与例 4.2.13 类似方法，可证明 $\lambda=0$ 不是 T 的特征值，故 $\lambda=0\in\sigma_c(T)$；归纳地，有

$$\mathbb{C}=\rho(T)\cup\sigma_p(T)\cup\sigma_c(T),$$

其中 $\sigma_p(T)=(0,1]$，$\sigma_c(T)=\{0\}$.

3. 谱集与正则集的性质

定理 4.2.9 设 $X\neq\{0\}$是非空 Banach 空间，$\mathfrak{B}X)$是 X 上的有界线性算子空间，对于 $T\in\mathfrak{B}X)$，有

(1) 算子 $T\in\mathfrak{B}X)$的谱集 $\sigma(T)\subset\mathbb{C}$是有界闭集；若 λ 是 T 的特征值，则对应于 λ 的全部特征向量与零向量，一起构成 X 的一个闭子空间，称为对应于 λ 的**特征向量空间**；若 $\lambda_k(k=1,2,\cdots,n)$是 T 的 n 个不同的特征值，x_k 是 T 的对应于λ_k 的任一个特征向量，则 x_1，x_2，$\cdots$，x_n 是线性无关的；

(2) 算子 $T\in\mathfrak{B}X)$的正则值集 $\rho(T)\subset\mathbb{C}$是开集；预解式 $R_\lambda=(\lambda I-T)^{-1}$作为 λ 的算子函数，在其定义域 $\rho(T)$中解析；

(3) 对于 $\lambda\in\mathbb{C}$，当 $|\lambda|>\|T\|$ 时，λ 是 T 的正则值，且 $\lambda I-T$ 的逆算子由公式

$$(\lambda I-T)^{-1}=\sum_{n=0}^{+\infty}\frac{1}{\lambda^{n+1}}T^n$$

给出，上式右边级数是按算子范数收敛，且 $\|(\lambda I-T)^{-1}\|\leqslant\frac{1}{|\lambda|-\|T\|}$；

(4) 对于 $\lambda\in\mathbb{C}$，若 λ 是 T 的正则值，则对于任何复数 $\mu\in\mathbb{C}$，当

$$|\mu-\lambda|<\|(\lambda I-T)^{-1}\|^{-1}$$

时，μ 也是正则值，且

$$(\mu I - T)^{-1} = \sum_{n=0}^{+\infty} (-1)^n (\mu - \lambda)^n (\lambda I - T)^{-(n+1)},$$

上式右边级数是按算子范数收敛.

请读者自行证明.

4. 可逆算子的性质

定理 4.2.10 设 X 是 Banach 空间，$\mathfrak{B}X)$ 是 X 上的有界线性算子空间，则

(1) $\mathfrak{B}X)$ 中可逆算子的全体是 $\mathfrak{B}X)$ 中的开集；

(2) 若 X 是含有非零元的 Banach 空间，则 $\forall\, T \in \mathfrak{B}X)$，必有 $\sigma(T) \neq \varnothing$.

5. 谱半径

定义 4.2.10（算子的谱半径） 设 $(X, +, \alpha\cdot, \|x\|_X)$ 是数域 $\mathbb{F}$ 上的 Banach 空间，对于任意有界线性算子 $T \in \mathfrak{B}X)$，称 $r_T = \max\limits_{\lambda \in \sigma(T)} |\lambda|$ 为算子 T 的**谱半径（spectral radius）**.

定理 4.2.11 设 X 是 Banach 空间，$\mathfrak{B}X)$ 是 X 上的有界线性算子空间，对于 $T \in \mathfrak{B}X)$，其谱半径为 $r_T = \lim\limits_{n \to +\infty} \sqrt[n]{\|T^n\|}$.

6. 紧算子及其性质

赋范空间上有一类线性算子非常重要，就是紧算子.

定义 4.2.11（紧算子） 设 $(X, \|x\|_X)$ 与 $(Y, \|y\|_Y)$ 是数域 $\mathbb{F}$ 上的 Banach 空间，对于线性算子 $T: X \to Y$，若 T 将 X 中的任一有界集 $E \subset X$ 映到 Y 中的序列紧集 $T(E) \subset Y$，则称 T 为**紧算子（compact operator）**. 紧算子也称为**全连续算子（totally continuous operator）**. X 到 Y 的紧算子的全体所成的集合，记为 $\mathfrak{C}(X, Y)$.

注 1 $T \in \mathfrak{C}(X, Y) \Leftrightarrow$ 对 X 中的任一有界集 $B \subset X$，集合 $\overline{T(B)}$ 是 Y 中的紧致集（即 $T(B)$ 有紧致闭包）.

注 2 $T \in \mathfrak{C}(X, Y) \Leftrightarrow$ 对 X 中的任一有界序列 $\{x_n\} \subset X$，序列 $\{T(x_n)\}$ 在 Y 中存在收敛子序列；亦即，$\{T(x_n)\}$ 是序列紧的.

定理 4.2.12（紧算子的连续性） 设 X 与 Y 是数域 $\mathbb{F}$ 上的 Banach 空间，若线性算子 $T: X \to Y$ 是紧算子，$T \in \mathfrak{C}(X, Y)$，则

(1) T 是连续算子，即有界算子；故 $\mathfrak{C}(X, Y) \subset \mathfrak{B}X, Y)$，并且 $\mathfrak{C}(X, Y)$ 在子拓扑之下是一个闭子空间；

(2) $T \in \mathfrak{B}X, Y)$ 是紧算子，当且仅当 $T: X \to Y$ 将 X 中的单位闭球 $S([0,1]) =$

$\{x\in X: \|x\|_X\leqslant 1\}$映成 X 中的序列紧集；

(3) $T\in\mathfrak{B}(X,Y)$是紧算子，当且仅当 $T:X\to Y$ 将 X 中的任一有界集 $E\subset X$ 映到 Y 中的紧致集 $\overline{T(E)}\subset Y$.

例 4.2.15 设 X 与 Y 是数域$\mathbb{F}$上的 Banach 空间，若 $T\in\mathfrak{B}(X,Y)$，且值域 $T(X)$是 Y 的有限维子空间，则 T 是紧算子 $T\in\mathfrak{C}(X,Y)$.

这是因为 T 是有界线性算子，故 T 将 X 中的有界集映到 $T(X)\subset Y$ 中的有界集，但 $T(X)$是有限维的，因此其中的有界集都是序列紧的，故 T 是紧算子.

例 4.2.16 设 $K(t,s)$在 $a\leqslant t\leqslant b, a\leqslant s\leqslant b$ 上连续，则由

$$(Tf)(t)=\int_a^b K(t,s)f(s)\mathrm{d}s$$

定义的算子 T 是 $C([a,b])$到 $C([a,b])$的紧算子，$K(t,s)$称为算子 T 的**核(kernel)**.

事实上，设 $A\subset C([a,b])$是有界集，则存在正数 $M>0$，使得 $\forall f\in A$，都有 $\|f\|_{C([a,b])}\leqslant M$. 显然，一方面，算子 $T(f)(t)$是 $C([a,b])$上的有界线性算子，故 $T(A)$是 $C([a,b])$中的一致有界集. 另一方面，我们有

$$\begin{aligned}|(Tf)(t_1)-(Tf)(t_2)| &\leqslant\int_a^b|K(t_1,s)-K(t_2,s)||f(s)|\mathrm{d}s\\ &\leqslant M\int_a^b|K(t_1,s)-K(t_2,s)|\mathrm{d}s,\end{aligned}$$

由 $K(t,s)$在正方形 $a\leqslant t\leqslant b, a\leqslant s\leqslant b$ 中的连续性，故 $\forall\varepsilon>0$，$\exists\delta>0$，使得对于任意 $t_1,t_2\in[a,b]$，只要 $|t_1-t_2|<\delta$，就有 $|K(t_1,s)-K(t_2,s)|<\dfrac{\varepsilon}{M(b-a)}$，$\forall s\in[a,b]$. 因此，对于所有的 $f\in A$，都有 $|(Tf)(t_1)-(Tf)(t_2)|<\varepsilon$. 这表明 $A\subset C([a,b])$的像集 $T(A)\subset C([a,b])$是等度连续的，故据 $C([a,b])$中集合序列紧性的定理 4.1.10，知 $T(A)$为序列紧的，从而 T 是紧算子.

例 4.2.17 设 $\sum\limits_{i,j=1}^{+\infty}|a_{ij}|^2<+\infty$，则由无穷矩阵 $T=[a_{ij}]_{1\leqslant i,j<+\infty}$决定一个由

$$l^2=\left\{x:x=(\xi_1,\xi_2,\cdots),\sum_{j=1}^{+\infty}|\xi_j|^2<+\infty\right\}$$

到 l^2 的线性算子 $y=T(x)$，使得 $T:x=(\xi_1,\xi_2,\cdots)\to y=(\eta_1,\eta_2,\cdots)$，其中 $\eta_j=\sum\limits_{i=1}^{+\infty}a_{ij}\xi_i\ (j=1,2,\cdots)$，则 $T:l^2\to l^2$是紧算子.

事实上，由 Cauchy 不等式，有

$$\begin{aligned}\|y\|_{l^2}^2=\sum_{j=1}^{+\infty}|\eta_j|^2=\sum_{j=1}^{+\infty}\left|\sum_{i=1}^{+\infty}a_{ij}\xi_i\right|^2 &\leqslant\sum_{j=1}^{+\infty}\left(\sum_{i=1}^{+\infty}|a_{ij}|^2\right)\left(\sum_{i=1}^{+\infty}|\xi_i|^2\right)\\ &=\left(\sum_{i,j=1}^{+\infty}|a_{ij}|^2\right)\|x\|_{l^2}^2,\end{aligned}$$

故 T 是 l^2 到 l^2 的有界线性算子. 进而可以证明 T 是紧算子.

设 $A\subset l^2$ 是任一有界集,故存在 $M>0$,使得对于所有 $x\in A$,有 $\|x\|_{l^2}<M$. 由于 $\sum_{i,j=1}^{+\infty}|a_{ij}|^2<+\infty$,故 $\forall\varepsilon>0$, $\exists N>0$, s. t. $\sum_{i=1}^{+\infty}\sum_{j=N+1}^{+\infty}|a_{ij}|^2<\frac{\varepsilon^2}{M^2}$. 从而

$$\sum_{j=N+1}^{+\infty}|\eta_j|^2=\sum_{j=N+1}^{+\infty}\left|\sum_{i=1}^{+\infty}a_{ij}\xi_i\right|^2\leqslant\sum_{j=N+1}^{+\infty}\left(\sum_{i=1}^{+\infty}|a_{ij}|^2\right)\left(\sum_{i=1}^{+\infty}|\xi_i|^2\right)\leqslant\frac{\varepsilon^2}{M^2}\|x\|_{l^2}^2\leqslant\varepsilon^2,$$

这表明算子 $T(x)=y$ 的前 N 个坐标 $\eta_1,\eta_2,\cdots,\eta_N$ 构成的元 $(\eta_1,\eta_2,\cdots,\eta_N,0,0,\cdots)$ 所成的集合

$$B=\{(\eta_1,\eta_2,\cdots,\eta_N,0,0,\cdots)\}$$

组成集合 $T(A)$ 的有限 $\varepsilon>0$ 网,而 B 是 l^2 中的 N 维子空间中的有界集,故它是序列紧的,这样,$T(A)$ 有序列紧的有限 ε 网,所以 $T(A)$ 是序列紧集,最后得到 T 是紧算子.

我们列出紧算子的一些性质而不加证明,供参考.

定理 4.2.13 (紧算子的复合与共轭) 设 X_1,X_2,X_3 是 Bnanch 空间,算子 $T\in\mathfrak{B}(X_1,X_2)$,$S\in\mathfrak{B}(X_2,X_3)$,若 T,S 中有一个是紧算子,例如 $T\in\mathfrak{C}(X_1,X_2)$,则复合 $S\circ T\in\mathfrak{C}(X_1,X_3)$ 也是紧算子;且共轭算子 T^* 也是紧算子,并成立 $T\in\mathfrak{C}(X_1,X_2)\Leftrightarrow T^*\in\mathfrak{C}(X_2^*,X_1^*)$.

定理 4.2.14(紧算子值域的可分性) 设 X,Y 是 Banach 空间,若 $T\in\mathfrak{C}(X,Y)$ 是紧算子,则 T 的值域 $\mathfrak{R}(T)\subset Y$ 在 Y 中是可分的.

定理 4.2.15(紧算子序列的极限定理) 设 X,Y 是 Banach 空间,$\{T_n\}\subset\mathfrak{C}(X,Y)$ 是紧算子序列,若对于算子 $T\in\mathfrak{B}(X,Y)$,有 $\lim\limits_{n\to\infty}\|T_n-T\|=0$,亦即,紧算子序列 $\{T_n\}$ 按范数收敛于有界算子 T,则 T 是紧算子,$T\in\mathfrak{C}(X,Y)$.

定理 4.2.16(紧算子的 $\lambda I-T$ 的性质) 设 X 是 Banach 空间,$T\in\mathfrak{C}(X)$ 是紧算子,我们有

(1) 若 $\lambda\neq0$,则算子 $\lambda I-T$ 的值域 $\mathfrak{R}(\lambda I-T)\subset X$ 是 X 中的闭子空间;

(2) 若 $\lambda\neq0$,则算子 $\lambda I-T$ 是满射,当且仅当 $\lambda I-T$ 为单射;因此,当 $\lambda I-T$ 为满射(或单射)时,λ 必为 T 的正则值.

于是,任一复数 $\lambda\neq0$ 是紧算子 T 的正则值 $\Leftrightarrow$ $\lambda I-T$ 是单射,或 $\lambda I-T$ 是满射.

定理 4.2.17(紧算子的谱定理) 设 X 是 Banach 空间,$T\in\mathfrak{C}(X)$ 是紧算子,则关于 T 的谱集有如下性质.

(1) $\sigma(T)\subset\mathbb{C}$ 或者是有限集,或者至多是以 $\lambda=0$ 为聚点的可数集;于是,$\sigma(T)$ 中任一不为零的数 $\lambda\in\sigma(T)$,$\lambda\neq0$ 都是 T 的特征值,亦即,$\sigma(T)\setminus\{0\}=\sigma_p(T)$;因此,$\sigma_p(T)$ 至多以 $\lambda=0$ 为聚点;

(2) 若 $\lambda\neq0$,$\lambda\in\sigma(T)$,$\lambda\in\sigma(T^*)$,则 T 与其共轭算子 T^* 对应于同一个 λ 的特征向量

空间有相同的维数；

(3) 当 X 为无限维时，必有 $0\in\sigma(T)$，$0\in\sigma(T^*)$；

(4) 对应于特征值 $\lambda\in\sigma_p(T)$ 的特征向量空间 $E_\lambda\subset X$ 是有限维的；

(5) T 与 T^* 对应于不同特征值 $\lambda_1,\lambda_1\in\sigma_p(T)$，$\lambda_1\neq\lambda_2$ 的特征向量空间 E_{λ_1} 与 E_{λ_2} 互相直交，$E_{\lambda_1}\perp E_{\lambda_2}$；

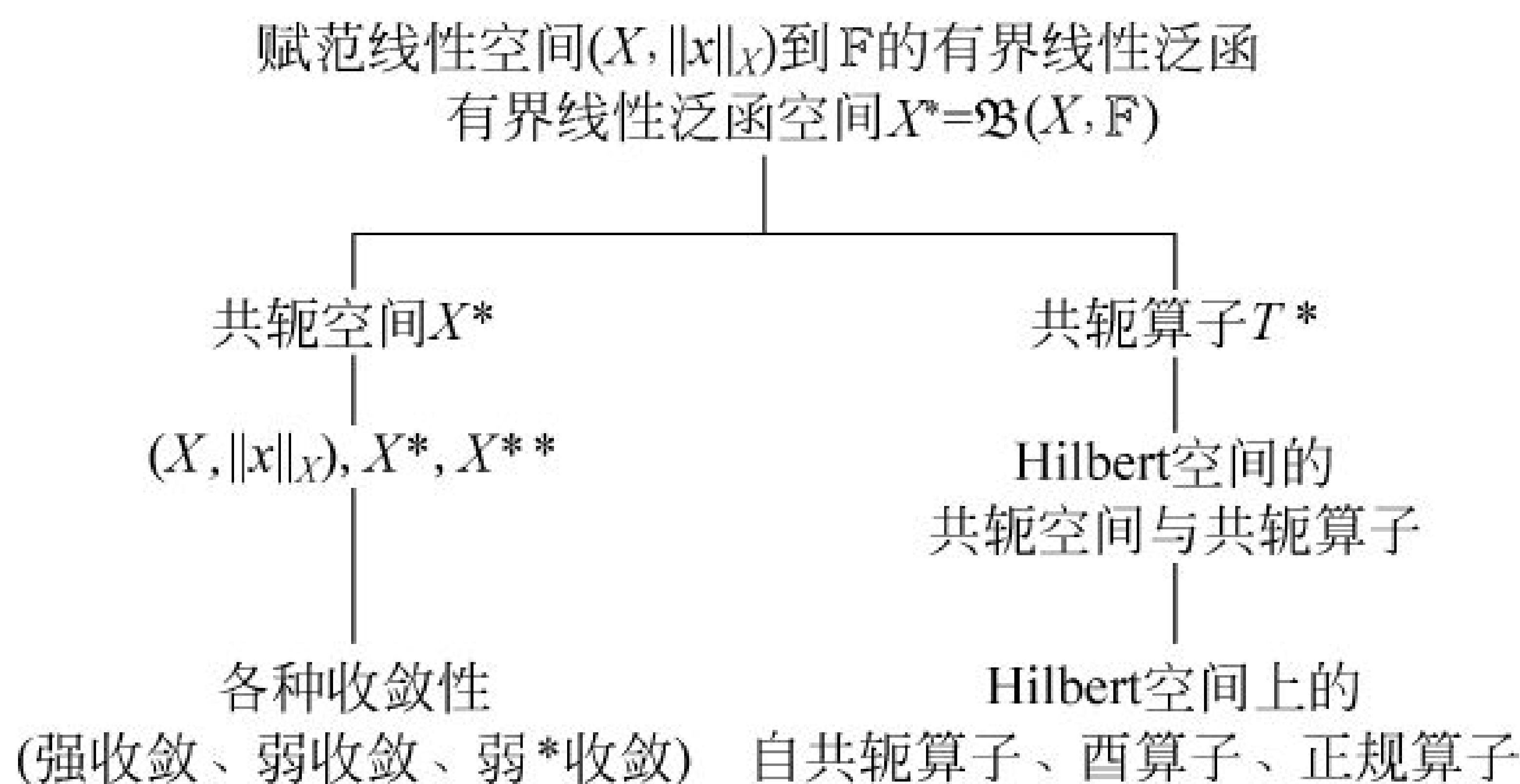
(6) 若复数 $\lambda\neq 0$ 是 T 的特征值，从而也是 T^* 的特征值，则方程 $(\lambda I-T)x=y$ 有解的充分必要条件是：y 与 λI^*-T^* 的零空间直交；方程 $(\lambda I^*-T^*)f=g$ 有解，当且仅当 g 与 $\lambda I-T$ 的零空间直交.

紧算子是一类非常重要的有界线性算子，有诸多应用，特别是在代数方程、算子方程、微分方程、积分方程等领域中.

4.3 线性泛函理论

4.3.1 赋范线性空间上的线性泛函

赋范线性空间 $(X,\|x\|_X)$ 上的有界线性泛函(连续线性泛函)的应用，不仅在数学科学自身的研究中随处可见，而且在自然科学各个领域中也居很重要的地位. 本节只选取其中一部分介绍，主要是有界线性泛函空间(共轭空间)，特别是其中的收敛性，共轭算子及其在 Hilbert 空间中的一些特殊算子，这些都是重要的基础.



1. 有界线性泛函

定义 4.3.1(线性泛函) 设 $(X,\|x\|_X)$ 是数域 $\mathbb{F}$ 上的赋范线性空间，称 $T:X\to\mathbb{F}$ 为**线性泛函(linear functional)**，若

$$T(x+y)=T(x)+T(y),\quad x,y\in X,$$
$$T(\alpha x)=\alpha T(x),\quad x\in X,\alpha\in\mathbb{F}.$$

若 $T\in\mathfrak{B}X,\mathbb{F})$，则称 T 为 X 上的**有界线性泛函(bounded linear functional)**，也称 T 为**连续**

线性泛函(continuous linear functional).

首先给出有界线性泛函的 Hahn-Banach 扩张定理.

定理 4.3.1 设$(X,\|x\|_X)$是赋范线性空间,$G\subset X$ 是 X 的子空间,G 上的有界线性泛函空间记为$(\mathfrak{B}(G,\mathbb{F}),+,\alpha\cdot,\|T\|_{\mathfrak{B}(G,\mathbb{F})})$. 若 $T\in\mathfrak{B}(G,\mathbb{F})$是 G 上的有界线性泛函,$T:G\to\mathbb{F}$,则 T 可保持范数不变地延拓到全空间 X 上,$\widetilde{T}:X\to\mathbb{F}$,$\widetilde{T}\in\mathfrak{B}(X,\mathbb{F})$;亦即,存在 X 上的有界线性泛函$\widetilde{T}:X\to\mathbb{F}$,使得

(1) $\widetilde{T}(x)=T(x),\forall x\in G$;

(2) $\|\widetilde{T}\|_{\mathfrak{B}(X,\mathbb{F})}=\|T\|_{\mathfrak{B}(G,\mathbb{F})}$.

注 1 Hahn-Banach 定理的证明非常精巧、细致,证明过程中用到 Zorn 引理,读者可以在很多泛函分析教程中看到.

注 2 由 Hahn-Banach 定理,今后可以不失一般性地假设有界线性泛函 $T\in\mathfrak{B}(X,\mathbb{F})$的定义域就是整个空间 X.

注 3 图 4.3.1 是 Hahn-Banach 扩张定理的示意图.

推论 1 设$(X,\|x\|_X)$是赋范线性空间,则$\forall x_0\in X$,$x_0\neq 0$,必存在 X 上的有界线性泛函 $f:X\to\mathbb{F}$,使得 f 的算子范数 $\|f\|=1$,且 $f(x_0)=\|x_0\|_X$.

推论 2 设$(X,\|x\|_X)$是赋范线性空间,$G\subset X$ 是子空间,$x_0\in X$,且若

$$d=\rho(x_0,G)=\inf_{x\in G}\|x-x_0\|_X>0,$$

则存在 X 上的有界线性泛函 $f:X\to\mathbb{F}$,其算子范数 $\|f\|$ 满足 $\|f\|=\dfrac{1}{d}$,且 $f(x)=\begin{cases}1, & x=x_0\\ 0, & x\in G.\end{cases}$

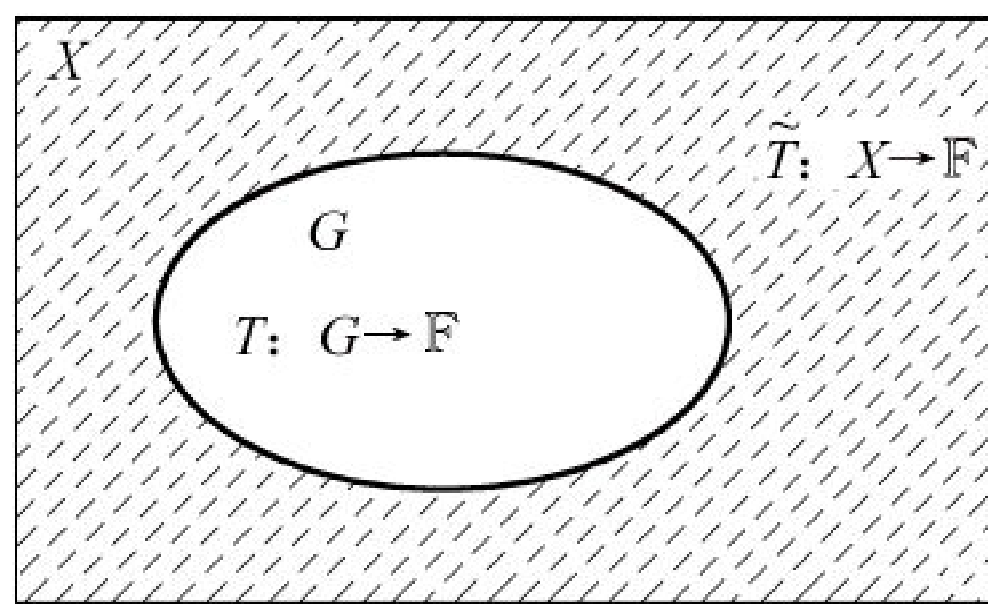


图 4.3.1 Hahn-Banach 扩张定理

例 4.3.1 取 $X=\mathbb{R}^2$,范数为 $\|x\|_X=|\xi_1|+|\xi_2|$(这是一个与 $\|x\|_{\mathbb{R}^2}=\sqrt{\xi_1^2+\xi_2^2}$ 等价的范数),$x=(\xi_1,\xi_2)\in\mathbb{R}^2$. 设 $G=\{x\in\mathbb{R}^2:x=(\xi_1,0)\}$,并定义 G 上的函数

$$f(x)=\xi_1,\quad x=(\xi_1,0)\in G.$$

显然,$f:G\to\mathbb{F}$ 是 G 上的有界线性泛函,且 $\|f\|=1$. 为将 f 扩张到整个空间$\mathbb{R}^2$,任取 $\alpha\in[-1,1]$,在$\mathbb{R}^2$ 上定义线性泛函$\tilde{f}:\mathbb{R}^2\to\mathbb{F}$,满足

$$\tilde{f}_\alpha(x)=\xi_1+\alpha\xi_2,\quad x=(\xi_1,\xi_2),$$

于是,$\tilde{f}_\alpha:\mathbb{R}^2\to\mathbb{F}$ 是 $f:G=\mathbb{R}^1\to\mathbb{F}$ 的延拓,且 $\|\tilde{f}_\alpha\|\geqslant\|f\|$;另一方面,

$$|\tilde{f}_\alpha(x)|\leqslant|\xi_1|+|\alpha\xi_2|\leqslant|\xi_1|+|\xi_2|=\|x\|_X,$$

由此算子范数为 $\|\tilde{f}_\alpha\|=\|f\|=1$.

2. 共轭空间与共轭算子

1）赋范线性空间的共轭空间

对于线性空间$(X,+,\alpha\cdot)$，我们定义了共轭空间(对偶空间，即线性空间上的线性泛函空间)，但那里没有涉及任何拓扑结构. 现在，要讨论赋范线性空间的共轭空间，即赋范线性空间上的线性泛函空间，并同时考虑其运算结构与拓扑结构.

定义 4.3.2(共轭空间) 设$X\equiv(X,+,\alpha\cdot,\|x\|_X)$是数域$\mathbb{F}$($\mathbb{R}$或$\mathbb{C}$)上的赋范线性空间，称$X$上有界线性泛函的全体组成的 Banach 空间$(\mathfrak{B}X,\mathbb{F}),+,\alpha\cdot,\|T\|)$为$X$的**共轭空间(conjugate space)**，或称**对偶空间**，记为$X^*=(\mathfrak{B}X,\mathbb{F}),+,\alpha\cdot,\|T\|)$，简记为$X^*=\mathfrak{B}X,\mathbb{F})$，其中范数为$\|T\|=\sup\left\{\dfrac{|T(x)|}{\|x\|_X}:x\in X,x\neq 0\right\}$.

我们也常将X的共轭空间X^*中的元记为$f\in X^*$，今后，用记号$\langle f,x\rangle\equiv f(x)\in\mathbb{F}$表示线性泛函$f\in X^*$对$x\in X$的作用.

对于X^*，可以再定义$(X^*)^*$，称为X的二次共轭，记为X^{**}. 还可以定义三次共轭、四次共轭等.

定义 4.3.3(自然嵌入映射) 设X是数域$\mathbb{F}$上的赋范线性空间，对于X^*与X^{**}，满足

$$\langle x^{**},f\rangle=\langle f,x\rangle,\quad \forall f\in X^*$$

的映射$\tau:X\to X^{**}$称为X到X^{**}的**自然嵌入映射(natural imbedding mapping)**(图 4.3.2).

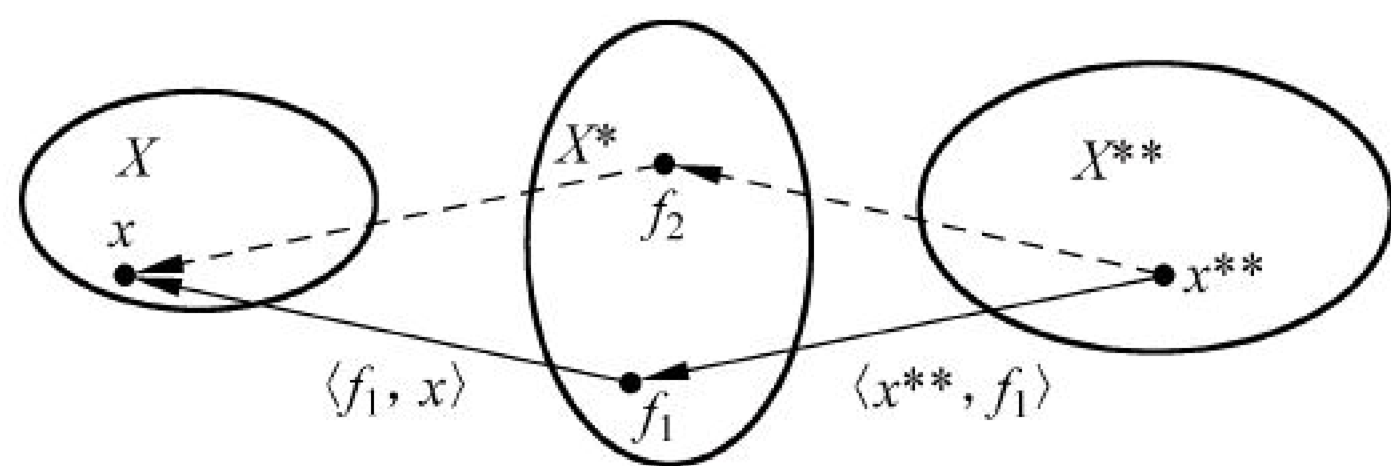


图 4.3.2 二次共轭、X到X^{**}的自然同构τ

设X是数域$\mathbb{F}$上的赋范线性空间，当$X^{**}=X$时，称X为**自反空间(self-reflect space)**. 当$X^*=X$时，称X为**自共轭空间(self-conjugate space)**.

对于自反空间$X^{**}=X$，自然嵌入映射$\tau:X\to X^{**}$是X到X^{**}的上等距同构映射：亦即，$\tau(x)=x^{**}$满足$\langle x^{**},f\rangle=\langle f,x\rangle,\forall f\in X^*$，且$\|x\|_X=\|x^{**}\|_{X^{**}}$.

2）几个重要空间的共轭空间

① $\mathbb{R}^n$的共轭空间是$\mathbb{R}^n$，$n\in\mathbb{Z}^+$. 亦即，$(\mathbb{R}^n)^*=\mathbb{R}^n$.

$$T\in(\mathbb{R}^n)^*\Rightarrow T=\xi=(\xi_1,\xi_2,\cdots,\xi_n)\in\mathbb{R}^n\Rightarrow T(x)=\langle\xi,x\rangle,\quad x\in\mathbb{R}^n,$$

其中$T(x)=\langle\xi,x\rangle=(\xi_1,\xi_2,\cdots,\xi_n)\cdot(x_1,x_2,\cdots,x_n)=\sum\limits_{j=1}^n\xi_jx_j$. 因此$\mathbb{R}^n$是自共轭的，并且

是自反的.

② $l^p(1<p<+\infty)$的共轭空间是 $l^q(1<q<+\infty)$,其中$\frac{1}{p}+\frac{1}{q}=1$,$q$ 称为 p 的共轭数. 亦即,$(l^p)^*=l^q$. 并且 $l^p(1<p<+\infty)$是自反的,即

$$(l^p)^{**}=l^p.$$

又 $T\in(l^p)^*=l^q \Rightarrow T=\xi=(\xi_1,\xi_2,\cdots,\xi_n,\cdots)\in l^q \Rightarrow T(x)=\langle\xi,x\rangle$ 满足

$$T(x)=\langle\xi,x\rangle=(\xi_1,\xi_2,\cdots,\xi_n,\cdots)\cdot(x_1,x_2,\cdots,x_n,\cdots)=\sum_{j=1}^{+\infty}\xi_j x_j,\quad \forall x\in l^p.$$

③ $L^p(\mathbb{R})(1<p<+\infty)$的共轭空间是 $L^q(\mathbb{R})(1<q<+\infty)$,其中$\frac{1}{p}+\frac{1}{q}=1$,$q$ 为 p 的共轭数. 亦即,$(L^p(\mathbb{R}))^*=L^q(\mathbb{R})$. 并且 p 幂可积函数空间是自反的,即

$$(L^p(\mathbb{R}))^{**}=L^p(\mathbb{R}).$$

又 $T\in(L^p(\mathbb{R}))^*(1<p<+\infty)$,$T\leftrightarrow g\in L^q(\mathbb{R})$ 满足

$$T(f)(g)=\langle f,g\rangle=\int_{\mathbb{R}}g(x)f(x)\mathrm{d}x,\quad \forall g\in L^p(\mathbb{R}).$$

④ 当 $p=q=2$ 时,有$(l^2)^*=l^2$;$(L^2(\mathbb{R}))^*=L^2(\mathbb{R})$,亦即,$l^2$,$L^2(\mathbb{R})$都是自反的、自共轭的 Hilbert 空间.

3）共轭算子

① 共轭算子的定义

定义 4.3.4(共轭算子) 设 X,Y 是数域$\mathbb{F}$上的赋范线性空间,X^*,Y^* 分别是它们的共轭空间. 对于 X 到 Y 的任一有界线性算子 $T\in\mathfrak{B}(X,Y)$,定义 T 的共轭算子,$T^*:Y^*\to X^*$ 为满足

$$\langle T^*(f),x\rangle=\langle f,T(x)\rangle,\quad \forall f\in Y^*,\quad \forall x\in X \tag{4.3.1}$$

的映射.

上述定义左边$\langle T^*(f),x\rangle$表示算子 $T^*\in\mathfrak{B}(Y^*,X^*)$的定义域是 Y^*,对于 Y^* 中的元 $f\in Y^*$ 的作用,记为 $T^*(f)$,于是得到 $T^*:f\in Y^*\xrightarrow{T^*\in\mathfrak{B}(Y^*,X^*)}T^*(f)\in X^*$;所以,$T^*(f)\in X^*=\mathfrak{B}(X,\mathbb{F})$;故 $T^*(f):X\to\mathbb{F}$,$\forall x\in X$ 的作用$\langle T^*(f),x\rangle$有意义,且$\langle T^*(f),x\rangle\in\mathbb{F}$.

另一方面,对于等式(4.3.1)的右边,由于 $f\in Y^*$,故 $f:Y\to\mathbb{F}$ 对 Y 中的元 $y\in Y$ 作用记为$\langle f,y\rangle$,$\forall f\in Y^*$,$\forall y\in Y$;于是,$\langle f,T(x)\rangle\in\mathbb{F}$,$\forall f\in Y^*$,$\forall x\in X$. 因此定义的共轭算子 $T^*:Y^*\to X^*$ 有意义. 回忆定义 2.2.3、(2.2.1)式及图 2.2.1,并作比较. 图 4.3.3 实际上是一个交换图.

② 共轭算子的性质

定理 4.3.2 设 X,Y,Z 为数域 $\mathbb{F}$ 上的赋范线性空间,则

(1) $S,T\in\mathfrak{B}(X,Y)\Rightarrow(S+T)^*=S^*+T^*$;

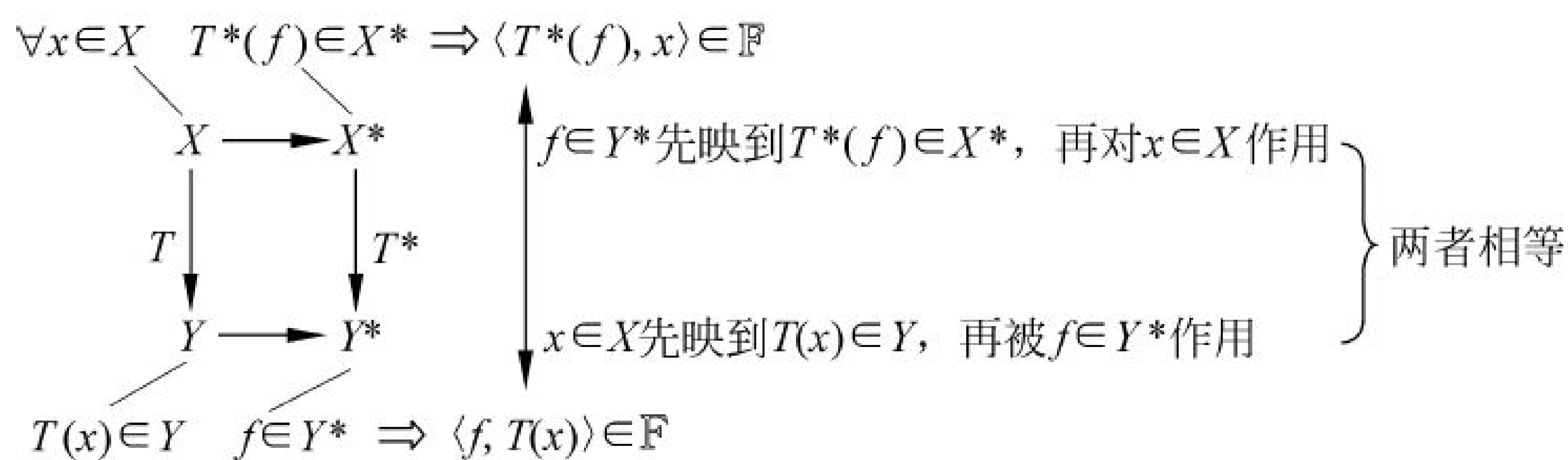


图 4.3.3 交换图

(2) $T\in\mathfrak{B}X,Y),\alpha\in\mathbb{F}\Rightarrow(\alpha T)^*=\alpha T^*$；

(3) $T\in\mathfrak{B}X,Y)\Rightarrow T^*\in\mathfrak{B}Y^*,X^*)$，且 $\|T^*\|=\|T\|$；

(4) $S\in\mathfrak{B}X,Y),T\in\mathfrak{B}Y,Z)\Rightarrow(T\circ S)^*=S^*\circ T^*$；

(5) 若 $T\in\mathfrak{B}X,Y)$的逆算子 T^{-1} 存在，$T^{-1}\in\mathfrak{B}Y,X)$，则共轭算子 $T^*\in\mathfrak{B}Y^*,X^*)$的逆算子$(T^*)^{-1}$也存在，且$(T^*)^{-1}=(T^{-1})^*$，并成立 $\|T^{-1}\|=\dfrac{1}{\|T\|}$，$\|(T^*)^{-1}\|=\dfrac{1}{\|T^*\|}$；

(6) 若 X,Y 为有限维，则 $T\in\mathfrak{B}X,Y)\leftrightarrow A=[T]_{\mathfrak{X}}$，矩阵 $A=[T]_{\mathfrak{X}}$称为算子 $T\in\mathfrak{B}X,Y)$的表示矩阵. 表示矩阵的秩 rank(A)定义为算子的秩 rank$(T)=$ rank(A). 我们有 rank$(T)=$ rank(T^*).

例 4.3.2 赋范线性空间 $X=\mathbb{R}^n$ 与 $Y=\mathbb{R}^m$，算子 $T\in L(\mathbb{R}^n,\mathbb{R}^m)\leftrightarrow A\in\mathfrak{M}_{m\times n}$的共轭算子是

$$T^*\in L(\mathbb{R}^m,\mathbb{R}^n)\leftrightarrow\overline{A}^{\mathrm{T}}\in\mathfrak{M}_{n\times m}.$$

因为$\mathbb{R}^n,\mathbb{R}^m$ 都是 Banach 空间，故算子 T,T^* 满足定理 4.3.2 中的结论.

例 4.3.3 若 T 是 Banach 空间 $X=L^p([a,b])(1<p<+\infty)$上以 $K(t,s)$为核的积子算子

$$(T(x))(t)=\int_{[a,b]}K(t,s)x(s)\mathrm{d}s,\quad x\in L^p([a,b]),\ t\in[a,b],$$

其中 $K(t,s)$是 $t,s\in[a,b]$的二元实值可测函数，满足 $\displaystyle\int_{[a,b]\times[a,b]}|K(t,s)|^q\mathrm{d}t\mathrm{d}s<+\infty\ (1<q<+\infty)$，$\dfrac{1}{p}+\dfrac{1}{q}=1$. 求 T^*.

解 我们断言：

(i) T 是 $L^p([a,b])$上的有界线性算子，且$(T(x))(t)\in L^q([a,b])$；故共轭算子 T^* 是 $L^q([a,b])$上的有界线性算子，且$\langle T^*(f),x\rangle=\langle f,T(x)\rangle$，$\forall f\in(L^q([a,b]))^*$，$\forall x\in L^p([a,b])$；

(ii) $\forall f\in(L^q([a,b]))^*$，$\exists y(t)\in L^p([a,b])$，s. t. $\langle f,z\rangle=\displaystyle\int_{[a,b]}y(t)z(t)\mathrm{d}t$，$\forall z\in L^q([a,b])$；

(iii) $(T^*(f))(s)=\int_{[a,b]}K(t,s)f(t)\mathrm{d}t, f\in L^p([a,b]), s\in[a,b]$,决定的算子 T^* 即为所求.

事实上,由(i)与(ii),T 的共轭算子 T^* 应满足:$\forall f\in(L^q([a,b]))^*$,有

$$\langle T^*(f),x\rangle=\langle f,T(x)\rangle=\left\langle f,\int_{[a,b]}K(t,s)x(s)\mathrm{d}s\right\rangle,\quad \forall x\in L^p([a,b]),$$

根据 Fubini 定理,

$$\langle T^*(f),x\rangle=\left\langle f,\int_{[a,b]}K(t,s)x(s)\mathrm{d}s\right\rangle=\int_{[a,b]}f(t)\left[\int_{[a,b]}K(t,s)x(s)\mathrm{d}s\right]\mathrm{d}t$$

$$=\int_{[a,b]}x(s)\left[\int_{[a,b]}K(t,s)f(t)\mathrm{d}t\right]\mathrm{d}s\equiv\left\langle\int_{[a,b]}K(t,s)f(t)\mathrm{d}t,x\right\rangle,$$

得到 $T^*(f)(s)=\int_{[a,b]}K(t,s)f(t)\mathrm{d}t,\forall f\in(L^q([a,b]))^*=L^p([a,b]),s\in[a,b]$.

③ 特殊的共轭算子

定义 4.3.5(特殊的共轭算子) 设 $T:X\to X$ 是数域 $\mathbb{F}$ 上的赋范线性空间 X 到自身的有界线性算子,$T\in\mathfrak{B}X)$的共轭算子为 $T^*\in\mathfrak{B}X^*)$.

(1) 若 $T=T^*$,则称 T 为**自共轭算子**,也称为 **Hermite 算子**;

(2) 若 T 为一一映射,且 $T^*=T^{-1}$,则称 T 为**酉算子**;

(3) 若 $TT^*=T^*T$,则称 T 为**正规算子**.

3. 收敛性

现在考虑 Banach 空间中的几种收敛性. 设 $X,Y,\cdots$是 Banach 空间,$X^*,Y^*,\cdots$是它们的共轭空间.

1) X 中的收敛性

设$\{x_n\}\subset X$ 为 Banach 空间 X 中的序列.

① 强收敛性

定义 4.3.6(强收敛性) 设$(X,\|x\|_X)$为 Banach 空间,$\{x_n\}\subset X$ 为 X 中的序列,若对于 $x\in X$,有

$$\lim_{n\to+\infty}\|x_n-x\|_X=0,$$

这里极限是在 X 范数意义下取的,则称**序列$\{x_n\}$强收敛到** x(或称**$\{x_n\}$按范数收敛到** x),称 x 为**$\{x_n\}$的强极限**,记为$(s)\lim\limits_{n\to+\infty}x_n=x$.

② 弱收敛性

定义 4.3.7(弱收敛性) 设$(X,\|x\|_X)$为 Banach 空间,$\{x_n\}\subset X$ 为 X 中的序列,若对

于 $x\in X$，有

$$\lim_{n\to+\infty}\langle f,x_n\rangle=\langle f,x\rangle,\quad \forall f\in X^*,$$

这里极限是在$\mathbb{R}$或$\mathbb{C}$范数意义下取的，则**称序列$\{x_n\}$弱收敛到**x，**称**x**是序列$\{x_n\}$的弱极限**，记为$(\mathrm{w})\lim\limits_{n\to+\infty}x_n=x$.

③ 收敛性的关系

定理 4.3.3 设$(X,\|x\|_X)$为 Banach 空间，则

(1) 强极限、弱极限都是惟一的；

(2) $(\mathrm{s})\lim\limits_{n\to+\infty}x_n=x\Rightarrow(\mathrm{w})\lim\limits_{n\to+\infty}x_n=x$；

(3) 若$\{x_n\}$弱收敛，则$\{\|x_n\|_X\}$是有界集；

(4) 若$(X,\|x\|_X)$是有限维的，则 $(\mathrm{s})\lim\limits_{n\to+\infty}x_n=x\Leftrightarrow(\mathrm{w})\lim\limits_{n\to+\infty}x_n=x$.

2) $\mathfrak{B}(X,Y)$中的收敛性

设$\{T_n\}\subset\mathfrak{B}(X,Y)$为有界线性算子空间$\mathfrak{B}(X,Y)$中的序列.

① 强收敛性

定义 4.3.8(强收敛性) 设$(X,\|x\|_X)$，$(Y,\|y\|_Y)$为 Banach 空间，$\{T_n\}\subset\mathfrak{B}(X,Y)$为有界线性算子序列，若对于$T\in\mathfrak{B}(X,Y)$，有

$$\lim_{n\to+\infty}\|T_n-T\|=0,$$

这里极限是在$\mathfrak{B}(X,Y)$范数意义下取的，则**称序列$\{T_n\}$强收敛到**T，或按$\mathfrak{B}(X,Y)$的算子范数$\|T\|$收敛到T，并**称**T**是序列$\{T_n\}$的强极限**，记为$(\mathrm{s})\lim\limits_{n\to+\infty}T_n=T$.

与定义 4.2.6 相比较，二者是一致的.

② 按点收敛性

定义 4.3.9(按点收敛性) 设$(X,\|x\|_X)$，$(Y,\|y\|_Y)$为 Banach 空间，$\{T_n\}\subset\mathfrak{B}(X,Y)$为有界线性算子序列，若对于$T\in\mathfrak{B}(X,Y)$，有

$$\lim_{n\to+\infty}T_n(x)=T(x),\quad\forall x\in X,$$

这里极限是在$(Y,\|y\|_Y)$的范数意义下取的，**称序列$\{T_n\}$按点收敛到**T(实际上这种收敛对于$(X,\|x\|_X)$中的每个点成立，故称“按点收敛”)；记为$(\mathrm{p})\lim\limits_{n\to+\infty}T_n=T$.

注意到，对于$(Y,\|y\|_Y)$而言，$T_n(x),T(x)\in Y$，按点收敛性的等价定义是

$$\lim_{n\to+\infty}T_n(x)=T(x),\ \forall x\in X\Leftrightarrow\lim_{n\to+\infty}\|T_n(x)-T(x)\|_Y=0,\ \forall x\in X,$$

故对于Y而言是“强”的. 与定义 4.2.7 相比较，二者是一致的.

③ 弱收敛性

定义 4.3.10(弱收敛性) 设$(X,\|x\|_X)$,$(Y,\|y\|_Y)$为 Banach 空间,$\{T_n\}\subset\mathfrak{B}(X,Y)$为有界线性算子序列,若对于$x\in X$,有

$$\lim_{n\to+\infty}\langle f,T_n(x)\rangle=\langle f,T(x)\rangle,\quad f\in Y^*,$$

其中$T_n(x)$,$T(x)\in Y$,$\forall x\in X$,极限是在$\mathbb{R}$或$\mathbb{C}$的范数意义下取的,则称**序列$\{T_n\}$弱收敛到T**;称T**是序列$\{T_n\}$的弱极限**,记为$(w)\lim\limits_{n\to+\infty}T_n=T$.

因此,在有界线性算子空间$\mathfrak{B}(X,Y)$中,有三种常用的收敛性.

④ 收敛性的关系

定理 4.3.4 设$(X,\|x\|_X)$,$(Y,\|y\|_Y)$是 Banach 空间,$\{T_n\}\subset\mathfrak{B}(X,Y)$,则

(1) 强极限、按点极限、弱极限都是惟一的;

(2) $(s)\lim\limits_{n\to+\infty}T_n=T\Rightarrow(p)\lim\limits_{n\to+\infty}T_n=T\Rightarrow(w)\lim\limits_{n\to+\infty}T_n=T$;

(3) 若$\{T_n\}$弱收敛,且$(X,\|x\|_X)$完备,则$\{\|T_n\|\}$是有界集;

(4) 若X,Y是有限维的,则

$$(s)\lim_{n\to+\infty}T_n=T\Leftrightarrow(p)\lim_{n\to+\infty}T_n=T\Leftrightarrow(w)\lim_{n\to+\infty}T_n=T.$$

图 4.3.4 直观地给出了各种收敛性的关系.

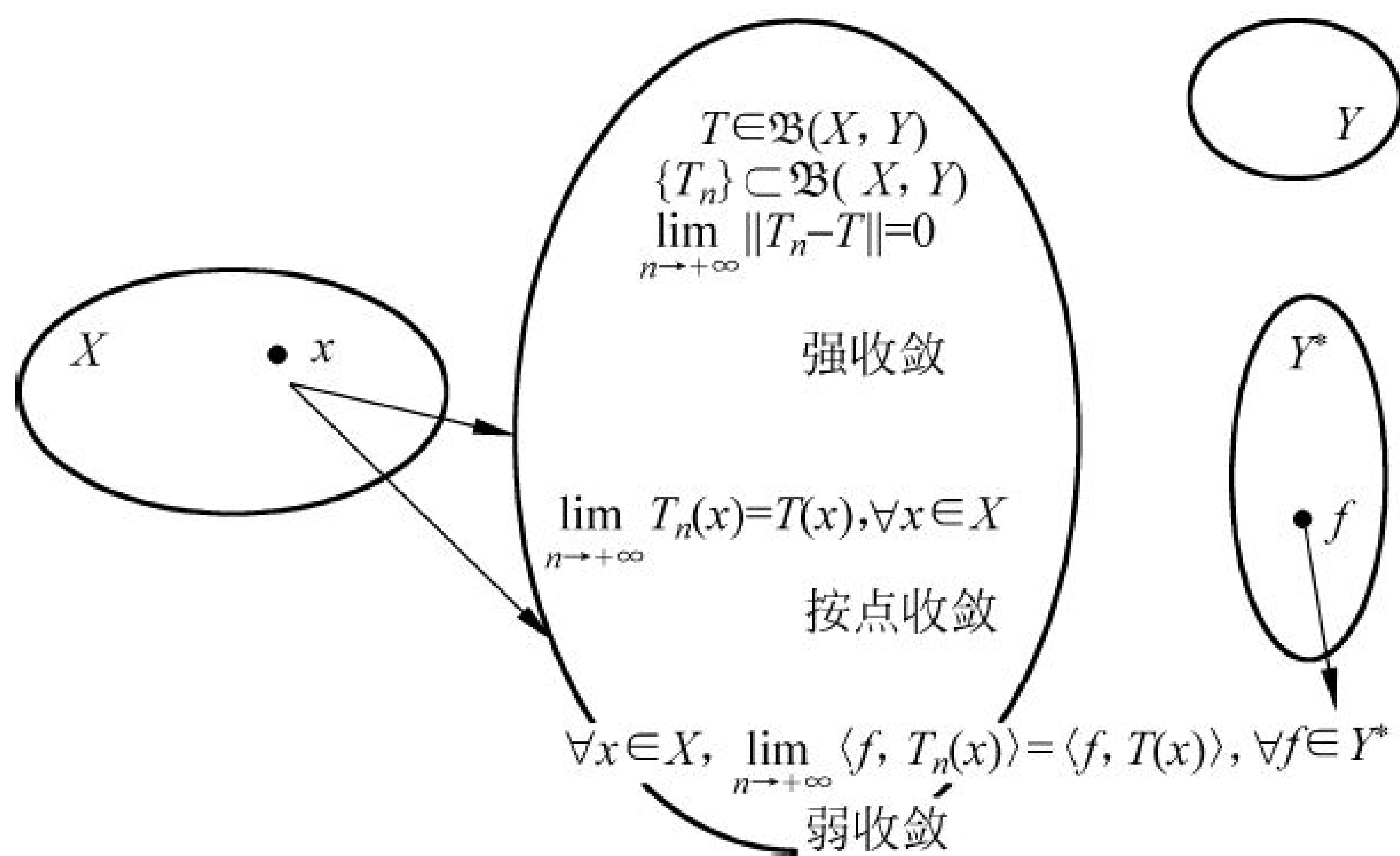


图 4.3.4 $\mathfrak{B}(X,Y)$中的三种收敛性

3) X^* 中的收敛性

设$\{f_n\}\subset X^*=\mathfrak{B}(X,\mathbb{F})$为有界线性泛函空间$X^*$中的序列.

① 强收敛性

定义 4.3.11(强收敛性) 设$(X,\|x\|_X)$,$(X^*,\|f\|_{X^*})$是 Banach 空间,$\{f_n\}\subset X^*$是有界线性泛函序列,若对于$f\in X^*$,有

$$\lim_{n\to+\infty}\|f_n-f\|_{X^*}=0,$$

这里极限是在 X^* 范数意义下取的，则**称序列$\{f_n\}$强收敛到** f，或称$\{f_n\}$按 X^* 的范数 $\|f\|_{X^*}$ 收敛到 f，**称 f 是序列$\{f_n\}$的强极限**，记为(s) $\lim\limits_{n\to+\infty} f_n=f$.

② 弱收敛性

定义 4.3.12(弱收敛性) 设$(X,\|x\|_X)$，$(X^*,\|x^*\|_{X^*})$，$(X^{**},\|x^{**}\|_{X^{**}})$分别为 Banach 空间、其一次共轭空间、二次共轭空间，序列$\{f_n\}\subset X^*$. 若对于 $x^{**}\in X$，有

$$\lim_{n\to+\infty}\langle x^{**},f_n\rangle=\langle x^{**},f\rangle,\qquad \forall x^{**}\in X^{**},$$

这里极限是在$\mathbb{R}$或$\mathbb{C}$范数意义下取的，**则称序列$\{f_n\}$弱收敛到** f，**称 f 是序列$\{f_n\}$的弱极限**，记为(w) $\lim\limits_{n\to+\infty} f_n=f$.

③ 弱*收敛性

定义 4.3.13(弱*收敛) 设$(X,\|x\|_X)$，$(X^*,\|x^*\|_{X^*})$分别为 Banach 空间、其一次共轭空间，序列$\{f_n\}\subset X^*$，若对于 $x\in X$，有

$$\lim_{n\to+\infty}\langle f_n,x\rangle=\langle f,x\rangle,\quad \forall x\in X,$$

这里极限是在$\mathbb{R}$或$\mathbb{C}$范数意义下取的，**则称序列$\{f_n\}$弱*收敛到** f；**称 f 是序列$\{f_n\}$的弱*极限**，记为 (w*) $\lim\limits_{n\to+\infty} f_n=f$.

因此，在有界线性泛函空间 $X^*=\mathfrak{B}(X,\mathbb{F})$中，也有三种常用的收敛性：强收敛性、弱收敛性、弱*收敛性.

④ 收敛性的关系

定理 4.3.5 设$(X,\|x\|_X)$是 Banach 空间，设$\{f_n\}\subset X^*$，则对于$\{f_n\}$的收敛性有如下性质：

(1) 强极限、弱极限、弱*极限都是惟一的；

(2) (s) $\lim\limits_{n\to+\infty} f_n=f\Rightarrow$ (w) $\lim\limits_{n\to+\infty} f_n=f\Rightarrow$ (w*) $\lim\limits_{n\to+\infty} f_n=f$；

(3) 若$\{f_n\}$弱*收敛，且$(X,\|x\|_X)$完备，则$\{\|f_n\|\}$是有界集；

(4) 若$(X,\|x\|_X)$是有限维的，则

$$(\mathrm{s})\lim_{n\to+\infty} f_n=f\Leftrightarrow(\mathrm{w})\lim_{n\to+\infty} f_n=f\Leftrightarrow(\mathrm{w}^*)\lim_{n\to+\infty} f_n=f;$$

(5) 若$(X,\|x\|_X)$是自反的，则

$$(\mathrm{w})\lim_{n\to+\infty} f_n=f\Leftrightarrow(\mathrm{w}^*)\lim_{n\to+\infty} f_n=f.$$

图 4.3.5 直观地给出了各种收敛性的关系.

4) 有限维赋范线性空间 X 的 X^*，X^{**} 中的收敛性

若$(X,\|x\|_X)$是有限维赋范线性空间，则 $X=X^*=X^{**}$，其中的各种收敛性以及其间的关系，请读者自行给出.

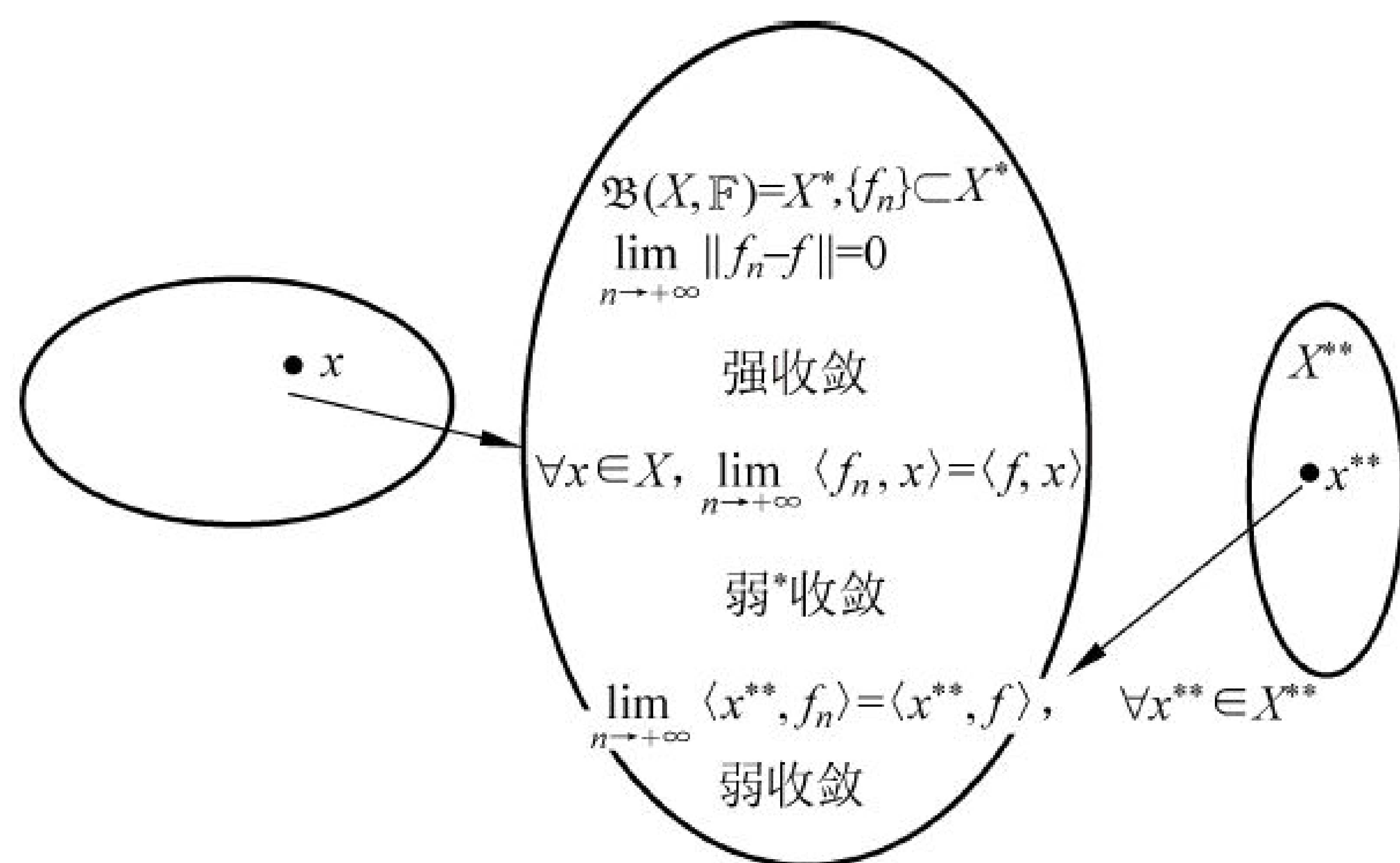


图 4.3.5 X^* 中的三种收敛性

例 4.3.4 $X=L^p([a,b]),1<p<+\infty$,序列$\{x_n\}\subset L^p([a,b])$收敛于 $x_0\in L^p([a,b])$.

(1) 作为 Banach 空间 $X=L^p([a,b])$.

① 强收敛(按范数收敛) (s) $\lim\limits_{n\to+\infty} x_n=x_0 \Leftrightarrow \lim\limits_{n\to+\infty}\|x_n-x_0\|_{L^p([a,b])}=0$

$$\Leftrightarrow \lim_{n\to+\infty}\left\{\int_{[a,b]}|x_n(t)-x_0(t)|^p\mathrm{d}t\right\}^{\frac{1}{p}}=0;$$

② 弱收敛 (w) $\lim\limits_{n\to+\infty} x_n=x_0 \Leftrightarrow \lim\limits_{n\to+\infty}\langle f,x_n\rangle=\langle f,x_0\rangle,\quad \forall f\in L^q([a,b])$

$$\Leftrightarrow \lim_{n\to+\infty}\int_{[a,b]}f(t)x_n(t)\mathrm{d}t=\int_{[a,b]}f(t)x_0(t)\mathrm{d}t,\quad \forall f\in L^q([a,b]).$$

(2) 作为 Banach 空间的共轭空间 $X^*=L^p([a,b]),\dfrac{1}{p}+\dfrac{1}{q}=1$.

①′ 强收敛(按范数收敛) (s) $\lim\limits_{n\to+\infty} x_n=x_0 \Leftrightarrow$ 与①相同;

②′ 弱收敛 (w) $\lim\limits_{n\to+\infty} x_n=x_0 \Leftrightarrow$ 由 $X^{**}=L^q([a,b])$,与②相同;

③′ 弱*收敛 (w*) $\lim\limits_{n\to+\infty} x_n=x_0 \Leftrightarrow$ 由 $X=L^q([a,b]),X^*=L^p([a,b])$,

$$\lim_{n\to+\infty}\langle x_n,x\rangle=\langle x_0,x\rangle,\quad \forall x\in L^q([a,b]),$$

$$\Leftrightarrow \lim_{n\to+\infty}\int_{[a,b]}x_n(t)x(t)\mathrm{d}t=\int_{[a,b]}x_0(t)x(t)\mathrm{d}t,\quad \forall x\in L^q([a,b]).$$

由于 $L^p([a,b])$是自反空间,因此弱收敛与弱*收敛相同.

例 4.3.5 $C([a,b])$中序列$\{x_n\}\subset C([a,b])$弱收敛于 $x_0\in C([a,b])$的一个充要条件:(w) $\lim\limits_{n\to+\infty} x_n=x_0 \Leftrightarrow$ (1) $\{\|x_n\|\}$为有界数列;(2) $\lim\limits_{n\to+\infty} x_n(t)=x_0(t)$处处成立.

例 4.3.6 $L^p([a,b]),1<p<+\infty$中的序列$\{x_n\}$弱收敛于 $x_0\in L^p([a,b])$的一个充要条件:(w) $\lim\limits_{n\to+\infty} x_n=x_0 \Leftrightarrow$ (1) $\{\|x_n\|\}$为有界数列;

$$(2)\ 对于每个\ c\in[a,b],有\ \lim_{n\to+\infty}\int_{[a,c]}x_n(t)\mathrm{d}t=\int_{[a,c]}x_0(t)\mathrm{d}t.$$

4.3.2 Hilbert 空间上的线性泛函

1. 共轭空间、共轭算子

设$(X,(x,y))$是数域 $\mathbb{F}$ 上的 Hilbert 空间，作为度量空间与赋范线性空间的特例，对于度量空间与赋范线性空间中的理论、定义、定理等都可以使用. 本节只考虑其作为 Hilbert 空间有哪些特别的性质.

1）Riesz 表现定理

有限维非奇异度量线性空间与内积空间上的 Riesz 表现定理已分别由定理 2.1.8 与定理 2.2.9 给出. 对于 Hilbert 空间，Riesz 表现定理有如下形式.

定理 4.3.6 设$(X,(x,y))$是 Hilbert 空间，则对 X 上的任一有界线性泛 $f\in X^*$，存在惟一的 $x\in X$，使得 $f(v)=(v,x)$，$\forall v\in X$，且 $\|f\|=\|x\|_X$，这里 $\|f\|$ 是 f 的算子范数.

2）Hilbert 空间的自共轭性、自反性

定理 4.3.7 设$(X,(x,y))$是 Hilbert 空间，则它是自共轭的自反空间，亦即，$X=X^*=X^{**}$.

事实上，由 Riesz 表现定理，可把 $f\in X^*$ 与 $x\in X$ 一一对应起来，$J:x\in X\leftrightarrow f\in X^*$，记为 $J(x)=f$，其中 $J:X\leftrightarrow X^*$ 是一一映射. 于是，对于 $x,y\in X$，$\forall v\in X$，有

$$\begin{aligned}\langle J(x+y),v\rangle &= (v,x+y) = (v,x)+(v,y)\\ &= \langle J(x),v\rangle+\langle J(y),v\rangle = \langle J(x)+J(y),v\rangle,\end{aligned}$$

故 $J(x+y)=J(x)+J(y)$. 又 $\forall v\in X$，$\forall \alpha\in\mathbb{F}$，有

$$\langle J(\alpha x),v\rangle = (v,\alpha x) = \bar{\alpha}(v,x) = \bar{\alpha}\langle J(x),v\rangle = \langle\bar{\alpha}J(x),v\rangle,$$

故 $J(\alpha x)=\bar{\alpha}J(x)$.

当 $\mathbb{F}=\mathbb{R}$ 时，$J:X\to X^*$ 是**线性算子**；当$\mathbb{F}=\mathbb{C}$ 时，$J:X\to X^*$ 是**共轭线性算子**，或其称为**反线性算子**.

进而，$\|J(x)\|_X=\|f\|=\|x\|_X$，故 $J:X\to X^*$ 是等距的. 于是，在等距同构意义下，X 是自共轭的，$X=X^*$，并且可以利用$(X,(x,y))$是 Hilbert 空间，也引进 X^* 上的内积. 记 $J^{-1}:X^*\to X$，则 $\forall f,g\in X^*\Rightarrow J^{-1}f,J^{-1}g\in X$. 令

$$(f,g)_{X^*} = (J^{-1}f,J^{-1}g)_X,$$

使得$(X^*,(f,g)_{X^*})$成为一个 Hilbert 空间. 这样，在等距同构意义下，X^* 也是自共轭空间，亦即 $X^*=(X^*)^*$. 从而得到 $X=X^*=X^{**}$.

3）共轭算子

设$(X,(x_1,x_2))$，$(Y,(y_1,y_2))$是两个实 Hilbert 空间，作为两个实线性空间，对于算子

$T\in L(X,Y)$，共轭算子 $T^*\in L(Y^*,X^*)$ 已在定义 2.2.3 给出；共轭算子 $T'\in L(Y,X)$ 也由定义 2.2.4 给出. 本小节讨论 Hilbert 空间 X 的有界线性算子空间 $\mathfrak{B}(X)\equiv\mathfrak{B}(X,X)$ 上的共轭算子.

定义 4.3.14（Hilbert 空间中的共轭算子） 设 $(X,(x,y))$ 是实 Hilbert 空间，$\mathfrak{B}(X)$ 是 X 到 X 的有界线性算子空间. 设 $T\in\mathfrak{B}(X)$，若算子 $T':X\to X$ 满足

$$(T'(y),x)=(y,T(x)),\quad \forall\, y\in X,\ \forall\, x\in X,$$

则称 T' 为 T 的**共轭算子**.

注 1 共轭算子 T' 是有界线性算子，$T'\in\mathfrak{B}(X)$.

注 2 对于两个 Hilbert 空间 X 与 Y，有界线性算子 $T\in\mathfrak{B}(X,Y)$ 的共轭算子 $T^*\in\mathfrak{B}(Y^*,X^*)$ 与 $T'\in\mathfrak{B}(Y,X)$ 之间，用类似于第 2 章中的证明方法，可以证明 T^* 与 T' 彼此拓扑同构. 对于特殊情形，在一个 Hilbert 空间 $(X,(x,y))$ 中，算子 $T\in\mathfrak{B}(X)$ 的共轭算子 $T^*\in\mathfrak{B}(X^*)=\mathfrak{B}(X)$ 与 $T'\in\mathfrak{B}(X)$ 彼此拓扑同构.

今后，在 Hilbert 空间 $(X,(x,y))$ 中，把 $T'\in\mathfrak{B}(X)$ 与 $T^*\in\mathfrak{B}(X^*)$ 视为等同，并用 T^* 记 T'，认定 $T^*\equiv T'\in\mathfrak{B}(X)$ 满足 $(T^*(y),x)=(y,T(x))$，$\forall\, y\in X$，$\forall\, x\in X$.

4）Hilbert 空间共轭算子的性质

定理 4.3.8 设 $(X,(x,y))$ 是实 Hilbert 空间，则 $T\in\mathfrak{B}(X)$ 的共轭算子 $T^*(\equiv T')\in\mathfrak{B}(X)$ 满足

（1）$T,S\in\mathfrak{B}(X)\Rightarrow(T+S)^*=T^*+S^*$；

（2）$T\in\mathfrak{B}(X),\alpha\in\mathbb{R}\Rightarrow(\alpha T)^*=\alpha T^*$；

（3）$T,S\in\mathfrak{B}(X)\Rightarrow(S\circ T)^*=T^*\circ S^*$；

（4）$T\in\mathfrak{B}(X)\Rightarrow\|T^*\|=\|T\|$；

（5）$T\in\mathfrak{B}(X)$，若逆算子 T^{-1} 存在，则 $(T^{-1})^*=(T^*)^{-1}$，且 $\|(T^{-1})^*\|=\dfrac{1}{\|T^*\|}$；

（6）$T\in\mathfrak{B}(X)\Rightarrow(T^*)^*=T$.

例 4.3.7 n 维欧氏空间 $X=\mathbb{R}^n$ 上，算子 $T\in\mathfrak{B}(\mathbb{R}^n)\leftrightarrow A\in\mathfrak{M}_{n\times n}$，设 $A=[\alpha_{ij}]_{n\times n}$，则算子

$$T:x=(\xi_1,\xi_2,\cdots,\xi_n)\in\mathbb{R}^n\to x'=(\xi_1',\xi_2',\cdots,\xi_n')\in\mathbb{R}^n,$$

由 $\xi_1'=\sum\limits_{j=1}^{n}\alpha_{kj}\xi_j$，$k=1,2,\cdots,n$ 确定. 求 T 的共轭算子 T^*：$\forall\, y=(\eta_1,\eta_2,\cdots,\eta_n)\in\mathbb{R}^n$，有

$$(x,T^*y)=(Tx,y)=\sum_{k=1}^{n}\left(\sum_{j=1}^{n}\alpha_{kj}\xi_j\right)\bar{\eta}_k=\sum_{j=1}^{n}\xi_j\overline{\sum_{k=1}^{n}\bar{\alpha}_{kj}\eta_k}=\overline{\sum_{k=1}^{n}\xi_k\sum_{j=1}^{n}\bar{\alpha}_{jk}\eta_j},$$

因此 T^* 是由 $(\bar{A})^{\mathrm{T}}=([\bar{\alpha}_{ji}]_{n\times n})^{\mathrm{T}}$ 决定的算子.

例 4.3.8 若 T 是 Hilbert 空间 $X=L^2([a,b])$上以 $K(t,s)$为核的积分算子，求 T^*.

$$(T(x))(t)=\int_{[a,b]}K(t,s)x(s)\mathrm{d}s,\quad x\in L^2([a,b]),\ t\in[a,b],$$

其中 $K(t,s)$是 $t,s\in[a,b]$的二元实值可测函数，满足 $\int_{[a,b]\times[a,b]}|K(t,s)|^2\mathrm{d}t\mathrm{d}s<+\infty$.

解 我们断言：

(i) T: $L^2([a,b])\to L^2([a,b])$是有界线性算子；

(ii) 由例 4.3.3 同样的推理，将$\langle\cdot,\cdot\rangle$换做内积$(\cdot,\cdot)$即可：

$$\exists\, y(t)\in L^2([a,b]),\text{使得 } f(z)=\int_{[a,b]}y(t)z(t)\mathrm{d}t,\ \forall z\in L^2([a,b])\ (\text{共轭算子定义})$$

$$\Rightarrow (x,T^*y)=(Tx,y)=\int_{[a,b]}\overline{y(t)}\left[\int_{[a,b]}K(t,s)x(s)\mathrm{d}s\right]\mathrm{d}t\ (\text{Fubini 定理})$$

$$=\int_{[a,b]}x(s)\overline{\left[\int_{[a,b]}\overline{K(t,s)}y(t)\mathrm{d}t\right]}\mathrm{d}s=\int_{[a,b]}x(t)\overline{\left[\int_{[a,b]}\overline{K(s,t)}y(s)\mathrm{d}s\right]}\mathrm{d}t$$

$$\Rightarrow (T^*y)(t)=\int_{[a,b]}\overline{K(s,t)}y(s)\mathrm{d}s\Rightarrow T^*\ \text{是以 } K^*(t,s)=\overline{K(s,t)}\ \text{为核的积分算子.}$$

例 4.3.9 求 Hilbert 空间 $X=L^2([0,1])$上的 Volterra 积分算子 T 的共轭算子 T^*：

$$(T(x))(t)=\int_{[0,t]}x(s)\mathrm{d}s,\quad x\in L^2([0,1]),\ t\in[0,1].$$

解 Volterra 积分算子 T 的核是 $K(t,s)=\begin{cases}1, & 0\leqslant t\leqslant 1, 0\leqslant s\leqslant t,\\ 0, & 0\leqslant t\leqslant 1, t<s\leqslant 1,\end{cases}$ 由例 4.3.8，共轭算子 T^* 的核为 $K^*(t,s)=\overline{K(s,t)}=\begin{cases}1, & 0\leqslant t\leqslant 1,\ t\leqslant s\leqslant 1,\\ 0, & 0\leqslant t\leqslant 1,\ 0\leqslant s<t,\end{cases}$ 于是，有

$$(T^*(x))(t)=\int_{[t,1]}x(s)\mathrm{d}s,\quad x\in L^2([0,1]),\ t\in[0,1].$$

2. Hilbert 空间中的自共轭算子、酉算子、正规算子

设$(X,(x,y))$是数域 $\mathbb{R}$ 上的 Hilbert 空间，有界算子空间为 $\mathfrak{B}(X)\equiv\mathfrak{B}(X,X)$.

定义 4.3.15（Hilbert 空间中的自共轭算子、酉算子、正规算子） 对于 $T\in\mathfrak{B}(X)$，

(1) 若 $T=T^*$，则称 T 为$(X,(x,y))$上的自共轭算子，亦称为 Hermite 算子；

(2) 若 T 为一一的，且 $T^*=T^{-1}$，则称 T 为$(X,(x,y))$上的酉算子；

(3) 若 $T^*T=TT^*$，则称 T 为$(X,(x,y))$上的正规算子.

例 4.3.10 Hilbert 空间 $X=L^2([0,1])$上的乘法算子 T：

$$(T(x))(t)=tx(t),\quad x\in L^2([0,1]),\ t\in[0,1].$$

解 乘法算子 T 是定义在 $L^2([0,1])$上、取值于 $L^2([0,1])$中的有界线性算子，因为线性性质是显然的.

由于 $\int_{[0,1]} |(T(x))(t)|^2 \mathrm{d}t = \int_{[0,1]} |tx(t)|^2 \mathrm{d}t \leqslant \int_{[0,1]} |x(t)|^2 \mathrm{d}t = \|x\|_{L^2([0,1])}^2 < +\infty$，故算子 $T: L^2([0,1]) \to L^2([0,1])$是有界算子，从而 $T \in \mathfrak{B}L^2([0,1]))$.

进而，(Tx,x)是 $L^2([0,1])$的内积，则有

$$(Tx,x) = \int_{[0,1]} T(x)(t) \cdot x(t)\mathrm{d}t = \int_{[0,1]} tx(t) \cdot x(t)\mathrm{d}t = \int_{[0,1]} x(t) \cdot tx(t)\mathrm{d}t = (x,Tx),$$

因此 $T: L^2([0,1]) \to L^2([0,1])$是 Hermite 算子.

方阵作为线性算子，有如下性质.

定理 4.3.9 设 $T \in \mathfrak{B}\mathbb{C})$为$\mathbb{C}^n \to \mathbb{C}^n$ 有界线性算子，在$\mathbb{C}^n$ 的给定基之下，$T \in \mathfrak{B}\mathbb{C})$与 $T^* \in \mathfrak{B}\mathbb{C})$分别有表示矩阵 $A \in \mathfrak{M}_n(\mathbb{C})$与 $B \in \mathfrak{B}_{\times n}(\mathbb{C})$，我们有

(1) 若 T 为自共轭算子，则 $A \in \mathfrak{M}_n(\mathbb{C})$为自共轭方阵，亦即$(\overline{A})^{\mathrm{T}} = A$；

(2) 若 T 为酉算子，则 $A \in \mathfrak{M}_n(\mathbb{C})$为酉方阵，亦即$(\overline{A})^{\mathrm{T}} = A^{-1}$；

(3) 若 T 为正规算子，则 $A \in \mathfrak{M}_n(\mathbb{C})$为正规方阵，亦即$(\overline{A})^{\mathrm{T}}A = A(\overline{A})^{\mathrm{T}}$.

进而，当 $T \in \mathfrak{B}\mathbb{R})$为$\mathbb{R}^n \to \mathbb{R}^n$ 有界线性算子时，有

(4) 若 T 为自共轭算子，则 $A \in \mathfrak{M}_n(\mathbb{R})$为对称方阵，亦即 $A^{\mathrm{T}} = A$；

(5) 若 T 为酉算子，则 $A \in \mathfrak{M}_n(\mathbb{R})$为正交方阵，亦即 $A^{\mathrm{T}} = A^{-1}$.

定理 4.3.10 (1) 设 $T \in \mathfrak{B}X)$，则 T 为 Hermite 算子，当且仅当

$$(Tx,y) = (x,Ty), \quad x,y \in X;$$

也当且仅当：对任何 $x \in X$，内积$(Tx,x) \in \mathbb{R}$；此时，称 Hermite 算子 T 为**对称算子**；

(2) 设 $T \in \mathfrak{B}X)$为 Hermite 算子，令 $m = \inf\limits_{\|x\|_X=1} (Tx,x)$，$M = \sup\limits_{\|x\|_X=1} (Tx,x)$，则算子 T 的范数为 $\|T\| = \max\{|m|,|M|\}$. 从而，Hermite 算子成立不等式

$$-\|T\|(x,x) \leqslant (Tx,x) \leqslant \|T\|(x,x);$$

(3) 设 $T \in \mathfrak{B}X)$为 Hermite 算子，则①T 的特征值都是实数；②对应于 T 的不同特征值的特征元(特征向量)互相直交；

(4) 设 $T \in \mathfrak{B}X)$为 Hermite 算子，则 T 的谱集 $\sigma(T) \subset \mathbb{R}$ 是有界闭集.

关于 Hilbert 空间上的三种重要算子的性质，可以类似定理 2.2.11～定理 2.2.13 得到. 还有许多重要性质及应用，读者可参考有关文献.

习题 4

1. 设(X,ρ)是度量空间，试证 $\rho_1(x,y) = \dfrac{\rho(x,y)}{1+\rho(x,y)}$也是 X 上的距离.

2. 在度量空间(X,ρ)中，将极限$\lim\limits_{n\to\infty} x_n = x_0$ 用 ε-δ 语言表示.

3. 试证：度量空间(X,ρ)中序列$\{x_n\}$，若极限存在，则此极限是惟一的.

4. 设 $f:X\to X_1$ 是由度量空间(X,ρ)到度量空间(X_1,ρ_1)的映射，试证：f 在点 $x_0\in X$ 连续，当且仅当对于任何收敛于 $x_0\in X$ 的序列$\{x_n\}\subset X$，有$\{f(x_n)\}$收敛到 $f(x_0)$.

5. 试证：度量空间(X,ρ)中任一收敛序列必定是 Cauchy 序列. 举出一个反例，说明此命题之逆命题不成立.

6. 度量空间(X,ρ)中两个 Cauchy 序列 $\xi=\{x_n\}$，$\eta=\{y_n\}$若满足 $\lim\limits_{n\to+\infty}\rho(x_n,y_n)=0$，则称 ξ 与 η"等价"，试证：这个"等价"$\xi=\{x_n\}\sim\eta=\{y_n\}$是一种等价关系.

7. 试证：度量空间中定义的 $\rho_0(\xi,\eta)=\lim\limits_{n\to+\infty}\rho(x_n,y_n)$满足(1)若 ξ,η 为同一等价类，则 $\rho_0(\xi,\eta)=0$；否则，此极限不为零；(2)$\rho_0(\xi,\eta)$满足距离的三点不等式.

8. 设 $X=\{x=(\xi_1,\xi_2,\cdots,\xi_k,0,0,\cdots):\xi_j\in\mathbb{R},1\leqslant j\leqslant k,k\in\mathbb{Z}^+\}$，试证：$X$ 是 l^p 的子空间，但在 l^p 的距离 $\rho(x,y)=\left(\sum\limits_{j=1}^{\infty}(x_j-y_j)^p\right)^{\frac{1}{p}}$ 之下，却不完备. 并且证明：l^p 是 X 的完备化.

9. 设 $X=\mathbb{Z}$ 是所有整数组成的集合，距离为 $d(m,n)=|m-n|$，试证其为完备度量空间.

10. 在实数全体组成的集合$\mathbb{R}$上，选定距离为 $d(x,y)=|\arctan x-\arctan y|$，试证：$(\mathbb{R},d)$是非完备的.

11. 试证：$C([a,b])$在距离 $\rho(f,g)=\left(\int_a^b|f(x)-g(x)|^p\mathrm{d}x\right)^{\frac{1}{p}}$ 之下不是完备度量空间，而 $L^p([a,b])$是 $C([a,b])$ 在距离上述 $\rho(f,g)$ 之下的完备化.

12. 试证：完备度量空间(X,ρ)的**闭球套性质**：若 $K_n=\overline{B(x_n,r_n)}\subset X$ 是 X 中的递减闭球序列，$K_1\supset K_2\supset\cdots\supset K_n\supset\cdots$，$n\geqslant1$，当球半径 r_n 所构成的序列满足 $\lim\limits_{n\to+\infty}r_n=0$ 时，则有惟一的点 $x_0\in X$，使得对于每个 $n\in\mathbb{Z}^+$，都有 $x_0\in\overline{B(x_n,r_n)}$.

13. 试证：在度量空间(X,ρ)中，(1)全有界子集 A 是有界的；(2)全有界子集 A 的有限 ε 网可取作 A 的有限子集 $B=\{x_1,x_2,\cdots,x_n\}\subset A$.

14. 试证：设$(X,\|x\|_X)$是局部紧致赋范线性空间，则单位球面 $S=\{x\in X:\|x\|_X=1\}$是紧致集.

15. 试证：两个度量空间之间的连续映射保持紧致性.

16. 试证：$\mathbb{N}$ 与 $\mathbb{Z}$ 不是$\mathbb{R}$的紧致子空间；单位闭区间$[0,1]$是$\mathbb{R}$中的一个紧致(子)空间.

17. 试证：在二维欧氏空间$\mathbb{R}^2$ 中，对于任一点 $x=(\xi_1,\xi_2)\in\mathbb{R}^2$，三种表示 $\|x\|_1=|\xi_1|+|\xi_2|$，$\|x\|_2=(|\xi_1|^2+|\xi_2|^2)^{\frac{1}{2}}$，$\|x\|_3=\max\{|\xi_1|,|\xi_2|\}$均是$\mathbb{R}^2$ 上的范数，并且三者是等价范数.

18. 设$(X,\|x\|)$为 n 维赋范线性空间，$\{e_1,e_2,\cdots,e_n\}$为 X 的基，对于 $x\in X$，令 $x=\sum\limits_{j=1}^{n}\xi_je_j$，$\xi_j\in\mathbb{F}$($\mathbb{F}$ 为$\mathbb{R}$ 或$\mathbb{C}$)，试证：X 到$\mathbb{F}^n$ 的映射 $T:T(x)=(\xi_1,\xi_2,\cdots,\xi_n)$是 X 到$\mathbb{F}^n$ 的同构映射.

19. 试证：若$(X,\|x\|)$是赋范线性空间，则$(X,\|x\|)$是局部紧致空间，当且仅当 X 中的任一有界闭集是紧致集(有时，也将此充要条件作为局部紧致空间的定义).

20. 试证：内积空间$(X,(x,y))$的内积(x,y)是 x 与 y 的连续函数.

21. 试证：在内积空间$(X,(x,y))$中，$x\perp y$ 的充要条件是：对所有实数 a，有 $\|x+ay\|=\|x-ay\|$，这里 $\|x\|$ 是由内积(x,y)决定的范数.

22. 在$\mathbb{R}^2$ 中，若(1)$M=\{x\in\mathbb{R}^2:x=(\xi_1,\xi_2)\neq0\}$，(2)$M$ 为$\mathbb{R}^2$ 中的线性无关集$\{x_1,x_2\}$，分别求出 $M^{\perp}$.

23. 设$(X,(x,y))$为 Hilbert 空间，$\{e_n\}\subset X$ 是 X 的标准直交系，则对于任一 $x\in X$，利用内积的性质，证明 Bessel 不等式 $\sum_{n=1}^{+\infty}|(x,e_n)|^2\leqslant\|x\|^2$.

24. 试证：算子范数有以下等价表示：$\|T\|=\sup\{\|T(x)\|_Y:\|x\|_X\leqslant 1,\forall x\in\mathfrak{D}\}$ 与 $\|T\|=\sup\{\|T(x)\|_Y:\|x\|_X=1,\forall x\in\mathfrak{D}\}$.

25. 设 $T:X\to Y$ 是赋范线性空间$(X,\|\cdot\|_X)$到赋范线性空间$(Y,\|\cdot\|_Y)$的线性算子，试证：T 在 $x=0\in X$ 连续，则 T 在整个 X 上连续.

26. 设 $C[a,b]$为$[a,b]$上的连续函数的全体. $f\in C[a,b]$的范数 $\|f\|=\max\limits_{x\in[a,b]}|f(x)|$使得 $C[a,b]$为 Banach 空间. 定义算子 $J(f)=\int_a^b f(x)\mathrm{d}x,\forall f\in C[a,b]$，试证：$J:C[a,b]\to\mathbb{R}$ 是 $C[a,b]$ 上的连续线性泛函.

27. 设$(X,\|x\|_X)$是赋范线性空间、$(Y,\|y\|_Y)$是 Banach 空间，试证：$(\mathfrak{B}(X,Y),\|T\|)$是 Banach 空间，这里的 $\|T\|$ 是算子范数.

28. 设$(\mathfrak{B}(X,Y),\|T\|)$是 Banach 空间$(X,\|x\|_X)$到 Banach 空间$(Y,\|y\|_Y)$的有界线性算子空间，子集族 $W\subset\mathfrak{B}(X,Y)$中的算子满足$\sup\limits_{T\in W}\|T(x)\|_Y<+\infty$，试证 $\|x\|_W$是 X 上的一个范数，其中 $\|x\|_W=\|x\|_X+\sup\limits_{T\in W}\|T(x)\|_Y$.

29. 设 X,Y 是赋范线性空间，$T:X\to Y$ 是闭线性算子，试证：(1)紧致子集 $C\subset X$ 的像集 $A=T(C)$在 X 中是闭的；(2)紧致子集 $K\subset X$ 的逆像集 $B=T^{-1}(K)$在 X 中是闭的.

30. 试证：算子一致收敛 $T_n\to T,T_n\in\mathfrak{B}(X,Y)$，蕴涵具有同一个极限 T 的算子的强收敛.

31. 赋范空间$(X,\|\cdot\|)$的子空间 $Y\subset X$ 在线性算子 $T:X\to Y$ 下，若 $T(Y)\subset Y$，则称 Y 为在映射 T 之下的不变子空间. 试证：T 的特征空间是 T 之下的不变子空间.

32. 设$\{e_k\}$为可分 Hilbert 空间$(H,(x,y))$中的完全标准直交序列，且 $T:H\to H$ 由 $Te_k=e_{k+1},k=1,2,\cdots$，定义. 将其线性与连续地扩张到整个 H 上，求其不变子空间并证明其无特征值.

33. 设 A 是含单位元的复 Banach 代数，若对于 $x\in A$，存在 $y,z\in A$，且 $yx=e,xz=e$，试证：x 为可逆的，且 $y=z=x^{-1}$.

34. 设$(H,(x,y))$是 Hilbert 空间，$T:H\to H$ 是有界线性双射算子，其逆 T^{-1}为有界的，试证：$(T^*)^{-1}$存在，且$(T^*)^{-1}=(T^{-1})^*$.

35. 设 S 与 T 为 Hilbert 空间$(H,(x,y))$上的线性算子，若在 H 上存在酉算子 U，使得 $S=UTU^{-1}=UTU^*$，则称算子 S 与 T 为酉等价. 若 T 为自共轭算子，试证：S 也是自共轭算子.

36. 试证：在定理 4.3.9 中，

(1) 若 T 为 Hermite 算子，则 $A\in\mathfrak{M}_n(\mathbb{C})$为 Hermite 方阵，亦即$(\bar{A})^{\mathrm{T}}=A$，且 $\|(\bar{A})^{\mathrm{T}}\|=\|A\|$；

(2) 若 T 为酉算子，则 $A\in\mathfrak{M}_n(\mathbb{C})$为酉方阵，亦即$(\bar{A})^{\mathrm{T}}=A^{-1}$，且 $\|(\bar{A})^{\mathrm{T}}\|=\|A^{-1}\|=\dfrac{1}{\|A\|}$；

(3) 若 T 为正规算子，则 $A\in\mathfrak{M}_n(\mathbb{C})$为正规方阵，亦即$(\bar{A})^{\mathrm{T}}A=A(\bar{A})^{\mathrm{T}}$.

进而，当 $T\in\mathfrak{B}(\mathbb{R}^n)$为$\mathbb{R}^n\to\mathbb{R}^n$ 有界线性算子时，有

(4) 若 T 为 Hermite 算子，则 $A\in\mathfrak{M}_n(\mathbb{R})$为对称方阵，亦即 $A^{\mathrm{T}}=A$，且 $\|A^{\mathrm{T}}\|=\|A\|$；

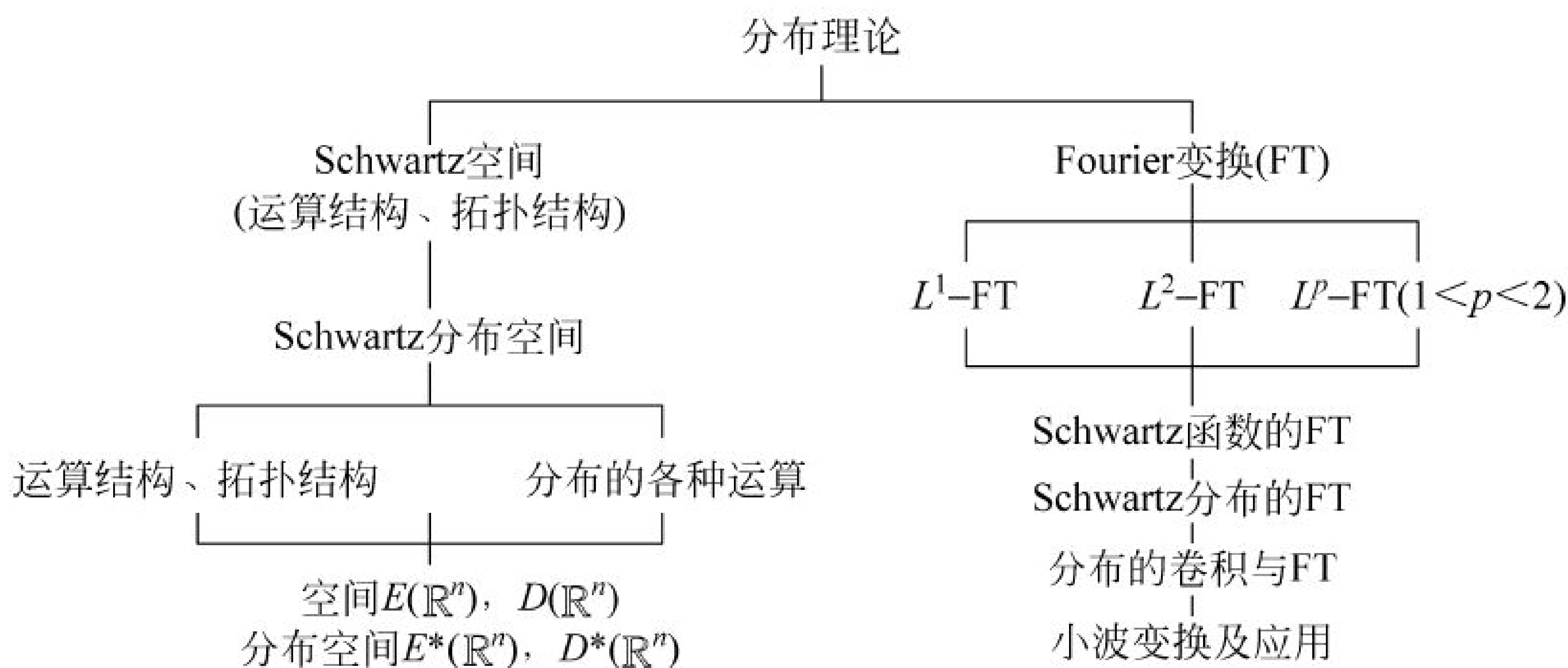
(5) 若 T 为酉算子，则 $A\in\mathfrak{M}_n(\mathbb{R})$为正交方阵，亦即 $A^{\mathrm{T}}=A^{-1}$，且 $\|A^{\mathrm{T}}\|=\|A^{-1}\|$.

第5章 分布理论

Fourier 分析对于各个科学领域的影响巨大，它是谱分析的基础，称其为"科学的基石(corner stone)"，当之无愧.

1822 年，法国数学家、物理学家 J. Fourier 发表了《热的解析理论》，把初始温度为正弦波的热传导方程的解，分解为具有"不同频率"的"谐波"，从而奠定了 Fourier 分析的基础. 在此后的近一个半世纪中，Fourier 分析在理论上得到不断完善，在应用中得到广泛推广，带动了数学科学、工程技术、医学科学等领域的蓬勃发展. 直到 20 世纪 50 年代，数学家完成了"分布理论"，使得 Fourier 分析臻于完美.

然而，数学科学的发展是无止境的，在 Fourier 分析的理论与应用都达到一个新的高度的时候，科学发展的新需求又促使这门学科生长出新的枝芽，例如"小波分析"、"分形分析"等，形成科学发展史上的靓丽风景.



本章主要参考文献是[4]，[5]，[8]，[10]，[11].

5.1 Schwartz 空间、Schwartz 分布空间

5.1.1 Schwartz 空间

1. 检验函数类

基本集合取为 n 维欧氏空间$\mathbb{R}^n=\{x=(x_1,x_2,\cdots,x_n):x_j\in\mathbb{R},j=1,2,\cdots,n\},n\in\mathbb{Z}^+$；

$\Omega\subseteq\mathbb{R}^n$ 为$\mathbb{R}^n$ 中的开集，$f:\Omega\to\mathbb{R}$ 为 Ω 上的实值函数. 令 $C^\infty(\Omega)=\{f:\Omega\to\mathbb{R}, f$ 是 Ω 上的任意次连续可导函数$\}$，$C_C^\infty(\Omega)=\{f\in C^\infty(\Omega): f$ 具有含在 Ω 中的紧致支集$\}$.

定义 5.1.1（检验函数类、Schwartz 函数类） 称

$$\mathbb{S}(\mathbb{R}^n)=\{\varphi\in C^\infty(\mathbb{R}^n):|x^\beta\partial^p\varphi(x)|\leqslant M_{\beta,p},\forall\beta,p\in\mathbb{N}^n\}$$

为 Schwartz 函数类（Schwartz function class），或检验函数类（Test function calss），其中 $M_{\beta,p}$ 是只与 β,p 有关的常数，$x^\beta\equiv x_1^{\beta_1}x_2^{\beta_2}\cdots x_n^{\beta_n}$，$x\in\mathbb{R}^n$，$\beta=(\beta_1,\beta_2,\cdots,\beta_n)\in\mathbb{N}^n$，$\partial^p\varphi(x)\equiv\dfrac{\partial^p\varphi(x)}{\partial x^p}\equiv\dfrac{\partial^{p_1+p_2+\cdots+p_n}\varphi(x)}{\partial x_1^{p_1}\partial x_2^{p_2}\cdots\partial x_n^{p_n}}$，$p=(p_1,p_2,\cdots,p_n)\in\mathbb{N}^n$.

2. Schwartz 空间

为使检验函数类$\mathbb{S}(\mathbb{R}^n)$成为一个拓扑线性空间，在$\mathbb{S}(\mathbb{R}^n)$中首先给出线性空间结构. 定义

$$\text{加法}\quad \varphi,\psi\in\mathbb{S}(\mathbb{R}^n)\Rightarrow(\varphi+\psi)(x)=\varphi(x)+\psi(x),\forall x\in\mathbb{R}^n;$$
$$\text{数乘}\quad \varphi\in\mathbb{S}(\mathbb{R}^n),\quad \alpha\in\mathbb{F}\Rightarrow(\alpha\varphi)(x)=\alpha\varphi(x),\forall x\in\mathbb{R}^n,$$

则$(\mathbb{S}(\mathbb{R}^n),+,\alpha\cdot)$成为**数域$\mathbb{F}$上的线性空间.**

对于线性空间$(\mathbb{S}(\mathbb{R}^n),+,\alpha\cdot)$，再定义“半范数族”，使其成为一个拓扑线性空间（参看拓扑线性空间的定义 3.4.13）.

称映射 $r:X\to[0,+\infty)$为 X 上的**半范数（semi-norm）**，若满足

(1) $r(\alpha x)=|\alpha|r(x),\forall x\in X,\alpha\in\mathbb{F}$，（绝对齐性）

(2) $r(x+y)\leqslant r(x)+r(y),\forall x,y\in X$.（次可加性）

（比较范数的定义，只相差 $r(x)=0\Leftrightarrow x=0$.）由半范数族可以确定 X 上的一个拓扑，使 X 成为一个拓扑空间.

$\forall\varphi\in\mathbb{S}(\mathbb{R}^n)$，定义映射 $r_{\beta,p}:\mathbb{S}(\mathbb{R}^n)\to[0,+\infty)$，使得

$$r_{\beta,p}(\varphi)=\sup_{x\in\mathbb{R}^n}|x^\beta\partial^p\varphi(x)|,\quad \beta,p\in\mathbb{N}^n,$$

则$\{r_{\beta,p}(\varphi):\beta,p\in\mathbb{N}^n\}$成为$\mathbb{S}(\mathbb{R}^n)$上的半范数族.

按照定义，可以证明$(\mathbb{S}(\mathbb{R}^n),+,\alpha\cdot,\{r_{\beta,p}(\varphi)\}_{\beta,p\in\mathbb{N}^n})$成为一个**拓扑线性空间.**

定理 5.1.1（$\mathbb{S}(\mathbb{R}^n)$的拓扑结构定理） $(\mathbb{S}(\mathbb{R}^n),+,\alpha\cdot,\{r_{\beta,p}(\varphi)\}_{\beta,p\in\mathbb{N}^n})$是一个具有可数半范数族的、$T_2$ 型的、完备的、可距离化的、自反的拓扑线性空间.

定理 5.1.2（$\mathbb{S}(\mathbb{R}^n)$的拓扑等价表示） $(\mathbb{S}(\mathbb{R}^n),+,\alpha\cdot,\{r_{\beta,p}(\varphi)\}_{\beta,p\in\mathbb{N}^n})$的拓扑结构可用该空间中“趋于 0 的序列”来定义，即$\forall\{\varphi_j\}\subset\mathbb{S}(\mathbb{R}^n)$，$\varphi_j\xrightarrow{\mathbb{S}(\mathbb{R}^n)}0$，当且仅当$\forall\beta,p\in\mathbb{N}^n$，$\lim\limits_{j\to+\infty}x^\beta\partial^p\varphi_j(x)=0$ 对于 $x\in\mathbb{R}^n$ 一致成立.

至此，我们又得到两种定义拓扑空间的方法：①由半范数族定义的拓扑；②给出“序列收敛于 0 的充要条件”. 当然还有其他的方法，例如归纳极限方法等.

例 5.1.1 $C_C^\infty(\mathbb{R}^n)\subset\mathbb{S}(\mathbb{R}^n)$.

例 5.1.2 $\varphi(x)=\mathrm{e}^{-\alpha|x|^2}\in\mathbb{S}(\mathbb{R}^n)$, $\alpha>0$.

5.1.2 Schwartz 分布空间

1. Schwartz 分布、Schwartz 分布空间

对 Schwartz 空间$\mathbb{S}(\mathbb{R}^n)\equiv(\mathbb{S}(\mathbb{R}^n),+,\alpha\cdot,\{r_{\beta,p}(\varphi)\}_{\beta,p\in\mathbb{N}^n})$,其共轭空间(对偶空间)

$$(\mathbb{S}(\mathbb{R}^n))^*\equiv\mathbb{S}^*(\mathbb{R}^n),$$

也就是$\mathbb{S}(\mathbb{R}^n)$上的连续线性泛函(连续性是指在$\mathbb{S}(\mathbb{R}^n)$的半范数拓扑之下)的全体所成的集合,称为 **Schwartz 分布空间(Schwartz distribution space)**,也称为**广义函数空间**;$\mathbb{S}^*(\mathbb{R}^n)$中的元 $T\in\mathbb{S}^*(\mathbb{R}^n)$称为 **Schwartz 分布(Schwartz distribution)**,也称为广义函数.

对于$\mathbb{S}^*(\mathbb{R}^n)\equiv(\mathbb{S}^*(\mathbb{R}^n),+,\alpha\cdot)$中的元,即 Schwartz 分布,$T\in\mathbb{S}^*(\mathbb{R}^n)$在每个元$\varphi\in\mathbb{S}(\mathbb{R}^n)$上的作用,记为$\langle T,\varphi\rangle=T(\varphi)$, $\forall\varphi\in\mathbb{S}(\mathbb{R}^n)$.

显然,在通常算子的运算下,$(\mathbb{S}^*(\mathbb{R}^n),+,\alpha\cdot)$成为数域$\mathbb{F}$上的**线性空间**:亦即,$\forall T_1$, $T_2\in\mathbb{S}^*(\mathbb{R}^n)$的线性运算 $\alpha_1T_1+\alpha_2T_2$ 定义为

$$\langle\alpha_1T_1+\alpha_2T_2,\varphi\rangle=\alpha_1\langle T_1,\varphi\rangle+\alpha_2\langle T_2,\varphi\rangle,\quad\forall\varphi\in\mathbb{S}(\mathbb{R}^n),\quad\alpha_1,\alpha_2\in\mathbb{F}.$$

2. Schwartz 分布的运算

对于 Schwartz 分布,可以定义以下运算.

(1) 分布的线性运算　分布 $T\in\mathbb{S}^*(\mathbb{R}^n)$的线性运算定义为

$$\langle T,\alpha_1\varphi_1+\alpha_2\varphi_2\rangle=\alpha_1\langle T,\varphi_1\rangle+\alpha_2\langle T,\varphi_2\rangle\quad\forall\varphi_1,\varphi_2\in\mathbb{S}(\mathbb{R}^n),\quad\alpha_1,\alpha_2\in\mathbb{F}.$$

亦即,每个分布 $T\in\mathbb{S}^*(\mathbb{R}^n)$是 Schwartz 空间$\mathbb{S}(\mathbb{R}^n)$上的**线性泛函**.

(2) 零分布　若分布 $T\in\mathbb{S}^*(\mathbb{R}^n)$满足

$$\langle T,\varphi\rangle=0,\quad\forall\varphi\in\mathbb{S}(\mathbb{R}^n),$$

则称此 $T\in\mathbb{S}^*(\mathbb{R}^n)$为**零分布(zero distribution)**,记为 0,$0\in\mathbb{S}^*(\mathbb{R}^n)$. 零分布就是线性空间$\mathbb{S}^*(\mathbb{R}^n)$中的加法的零元.

(3) 分布的相等　若分布 $T_1,T_2\in\mathbb{S}^*(\mathbb{R}^n)$满足

$$\langle T_1-T_2,\varphi\rangle=0,\quad\forall\varphi\in\mathbb{S}(\mathbb{R}^n),$$

亦即,T_1-T_2 为零分布,则称**分布 T_1 与 T_2 相等**,记为 $T_1=T_2$;

(4) 分布的平移　对于分布 $T\in\mathbb{S}^*(\mathbb{R}^n)$与 $a=(a_1,a_2,\cdots,a_n)\in\mathbb{R}^n$,定义满足

$$\langle U,\varphi\rangle=\langle T,\varphi_{-\boldsymbol{a}}\rangle,\quad\forall\varphi\in\mathbb{S}(\mathbb{R}^n)$$

的分布 $U\in\mathbb{S}^*(\mathbb{R}^n)$为**分布 $T\in\mathbb{S}^*(\mathbb{R}^n)$的平移(translation)**,其中 $\varphi_{-a}(x)=\varphi(x+a)$, $\forall\varphi\in\mathbb{S}(\mathbb{R}^n)$, $x\in\mathbb{R}^n$,并将分布 $T\in\mathbb{S}^*(\mathbb{R}^n)$的平移 $U\in\mathbb{S}^*(\mathbb{R}^n)$记为 T_a,或 τ_aT.

于是,分布 $T\in\mathbb{S}^*(\mathbb{R}^n)$的平移是一个满足

$$\langle T_{\boldsymbol{a}},\varphi\rangle=\langle T,\varphi_{-\boldsymbol{a}}\rangle,\quad\forall\varphi\in\mathbb{S}(\mathbb{R}^n)$$

的分布 $\tau_{\boldsymbol{a}}T=T_{\boldsymbol{a}}$.

(5) **分布的反射** 对于分布 $T\in\mathbb{S}^*(\mathbb{R}^n)$,定义满足

$$\langle V,\varphi\rangle=\langle T,\tilde{\varphi}\rangle,\quad \forall\varphi\in\mathbb{S}(\mathbb{R}^n)$$

的分布 $V\in\mathbb{S}^*(\mathbb{R}^n)$为**分布** $T\in\mathbb{S}^*(\mathbb{R}^n)$**的反射(reflection)**,其中 $\tilde{\varphi}(x)=\varphi(-x)$,且$(-x)=-(x_1,x_2,\cdots,x_n)=(-x_1,-x_2,\cdots,-x_n)$.将分布 $T\in\mathbb{S}^*(\mathbb{R}^n)$的反射 $V\in\mathbb{S}^*(\mathbb{R}^n)$记为$\tilde{T}$.

于是,分布 $T\in\mathbb{S}^*(\mathbb{R}^n)$的反射是一个满足

$$\langle\tilde{T},\varphi\rangle=\langle T,\tilde{\varphi}\rangle,\quad \forall\varphi\in\mathbb{S}(\mathbb{R}^n)$$

的分布$\tilde{T}$.

(6) **分布与函数的乘积** 对于分布 $T\in\mathbb{S}^*(\mathbb{R}^n)$与函数 $f\in C^\infty(\mathbb{R}^n)$,若 $f\varphi\in\mathbb{S}(\mathbb{R}^n)$;定义满足

$$\langle W,\varphi\rangle=\langle T,f\varphi\rangle,\quad \forall\varphi\in\mathbb{S}(\mathbb{R}^n)$$

的分布 $W\in\mathbb{S}^*(\mathbb{R}^n)$为**分布** $T\in\mathbb{S}^*(\mathbb{R}^n)$**与函数** $f\in C^\infty(\mathbb{R}^n)$**的乘积**,并将分布 $T\in\mathbb{S}^*(\mathbb{R}^n)$与 $f\in C^\infty(\mathbb{R}^n)$的乘积 $W\in\mathbb{S}^*(\mathbb{R}^n)$记为 fT.

于是,分布 $T\in\mathbb{S}^*(\mathbb{R}^n)$与函数 $f\in C^\infty(\mathbb{R}^n)$的乘积是一个满足

$$\langle fT,\varphi\rangle=\langle T,f\varphi\rangle,\quad \forall\varphi\in\mathbb{S}(\mathbb{R}^n)$$

的分布 fT.

(7) **分布的导数** 对于分布 $T\in\mathbb{S}^*(\mathbb{R}^n)$,定义满足

$$\langle Z,\varphi\rangle=-\langle T,\partial_{x_1}\varphi\rangle,\quad \forall\varphi\in\mathbb{S}(\mathbb{R}^n),$$

($\partial_{x_1}\varphi$ 是 φ 关于 x_1 的偏导数)的**分布** $Z\in\mathbb{S}^*(\mathbb{R}^n)$**为分布** $T\in\mathbb{S}^*(\mathbb{R}^n)$**关于** x_1 **的偏导数**,并将分布 $T\in\mathbb{S}^*(\mathbb{R}^n)$的偏导数记为$\partial_{x_1}T$.

于是,分布 $T\in\mathbb{S}^*(\mathbb{R}^n)$关于 x_1 的偏导数是一个满足

$$\langle\partial_{x_1}T,\varphi\rangle=-\langle T,\partial_{x_1}\varphi\rangle,\quad \forall\varphi\in\mathbb{S}(\mathbb{R}^n)$$

的分布$\partial_{x_1}T$. 类似地,可定义分布的高阶偏导数 $\dfrac{\partial^p T}{\partial x^p}\equiv\dfrac{\partial^{p_1+\cdots+p_n}T}{\partial x_1^{p_1}\cdots\partial x_n^{p_n}}$,$p=(p_1,\cdots,p_n)\in\mathbb{N}^n$.

(8) **分布的支集** 回顾函数的支集:$\forall\varphi(x)=\varphi(x_1,x_2,\cdots,x_n)$,$x=(x_1,x_2,\cdots,x_n)\in\mathbb{R}^n$,为定义在$\mathbb{R}^n$ 上的函数,定义 $\operatorname{supp}\varphi=\overline{\{x\in\mathbb{R}^n:\varphi(x)\neq0\}}$.

定义 5.1.2(分布的支集) 对于分布 $T\in\mathbb{S}^*(\mathbb{R}^n)$,若

$$\langle T,\varphi\rangle=0,\quad \forall\varphi\in\mathbb{S}(\mathbb{R}^n),\quad \operatorname{supp}\varphi\subset\Omega,$$

记为$T|_\Omega=0$,则**称分布** $T\in\mathbb{S}^*(\mathbb{R}^n)$**在开集** $\Omega\subset\mathbb{R}^n$ **上等于零**. 记$\{\Omega_\alpha\}_{\alpha\in I}$是使得分布 $T\in\mathbb{S}^*(\mathbb{R}^n)$为零所有开集所成的集合,则**并集** $\bigcup_{\alpha\in I}\Omega_\alpha$ **是使得分布** $T\in\mathbb{S}^*(\mathbb{R}^n)$**为零的最大开集**;**称并集** $\bigcup_{\alpha\in I}\Omega_\alpha$ **的补集**$\mathbb{R}^n\setminus\bigcup_{\alpha\in I}\Omega_\alpha$ **为分布** T **的支集**,记为$\operatorname{supp}T$,即

$$\operatorname{supp}T=\complement\{x\in\mathbb{R}^n:\exists\text{ 开集 }U,\text{使得}T|_U=0,x\in U\}.$$

易见,① $\operatorname{supp}T$ 为闭集,且 $\operatorname{supp}T=\overline{\{x\in\mathbb{R}^n:T\neq0\text{ 在 }x\text{ 的开邻域中成立}\}}$;

② 若 $T\in\mathbb{S}^*(\mathbb{R})$,$\varphi\in\mathbb{S}(\mathbb{R})$,且$\operatorname{supp}T\cap\operatorname{supp}\varphi=\varnothing$,则$\langle T,\varphi\rangle=0$.

3. Schwartz 分布空间的拓扑

赋予线性空间$(\mathbb{S}^*(\mathbb{R}^n),+,\alpha\cdot)$以拓扑结构，使其成为拓扑线性空间，一般有三种.

定义 5.1.3（Schwartz 分布空间的强拓扑） 赋予线性空间$(\mathbb{S}^*(\mathbb{R}^n),+,\alpha\cdot)$以算子范数拓扑$(\mathbb{S}^*(\mathbb{R}^n),+,\alpha\cdot,\|T\|_{\mathbb{S}^*(\mathbb{R}^n)})$，使其成为拓扑线性空间，称为$\mathbb{S}^*(\mathbb{R}^n)$上的强拓扑，或范数拓扑，记为$(\mathbb{S}^*(\mathbb{R}^n),+,\alpha\cdot,\tau_s)$.

定义 5.1.4（Schwartz 分布空间的弱拓扑） 对于连续线性泛函序列$\{T_k\}\subset\mathbb{S}^*(\mathbb{R}^n)$，定义序列$T_k\to 0(k\to+\infty)$为

$$T_k\xrightarrow{\tau_w}0(k\to+\infty)\Leftrightarrow\langle\psi,x^\beta\partial^p T_k\rangle\to 0(k\to+\infty),\quad\forall\psi\in\mathbb{S}^{**}(\mathbb{R}^n),\forall\beta,p\in\mathbb{N}^n,$$

这里$x^\beta\partial^p T_k$是 Schwartz 分布T_k的p次导数与单项式x^β的乘积，由此确定的拓扑称为$\mathbb{S}^*(\mathbb{R}^n)$上的弱拓扑，记为$(\mathbb{S}^*(\mathbb{R}^n),+,\alpha\cdot,\tau_w)$.

定义 5.1.5（Schwartz 分布空间的弱*拓扑） 对于连续线性泛函序列$\{T_k\}\subset\mathbb{S}^*(\mathbb{R}^n)$，定义序列$T_k\to 0(k\to+\infty)$为

$$T_k\xrightarrow{\tau_{w^*}}0(k\to+\infty)\Leftrightarrow\langle x^\beta\partial^p T_k,\varphi\rangle\to 0(k\to+\infty),\quad\forall\varphi\in\mathbb{S}(\mathbb{R}^n),\forall\beta,p\in\mathbb{N}^n,$$

由此确定的拓扑称为$\mathbb{S}^*(\mathbb{R}^n)$上的弱*拓扑，记为$(\mathbb{S}^*(\mathbb{R}^n),+,\alpha\cdot,\tau_{w^*})$.

定理 5.1.3 对于$(\mathbb{S}^*(\mathbb{R}^n),+,\alpha\cdot)$的拓扑结构，我们有

(1) 强拓扑等价于：对于$\mathbb{S}(\mathbb{R}^n)$中的任一有界集$A\subset\mathbb{S}(\mathbb{R}^n)$，

$$\langle x^\beta\partial^p T_k,\varphi\rangle\to 0(k\to+\infty),\quad\forall\beta,p\in\mathbb{N}^n,\forall\varphi\in A\text{ 一致成立；}$$

(2) 弱*拓扑与弱拓扑彼此等价；

(3) 在弱*拓扑之下，成立

$$\forall\varphi_k\xrightarrow{\mathbb{S}(\mathbb{R}^n)}0(k\to+\infty)\Rightarrow\langle T,x^\beta\partial^p\varphi_k\rangle\to 0(k\to+\infty),\quad\forall T\in\mathbb{S}^*(\mathbb{R}^n),\forall\beta,p\in\mathbb{N}^n.$$

因为$(\mathbb{S}(\mathbb{R}^n),+,\alpha\cdot,\{r_{\beta,p}(\varphi)\}_{\beta,p\in\mathbb{N}^n})$是自反空间，因此$(\mathbb{S}^*(\mathbb{R}^n),+,\alpha\cdot)$的弱*拓扑与弱拓扑一致.

4. Schwartz 分布的例子

例 5.1.3 $L^1(\mathbb{R}^n)\subset\mathbb{S}^*(\mathbb{R}^n)$.

$\forall f\in L^1(\mathbb{R}^n)$，$f$视为一个分布，在$\forall\varphi\in\mathbb{S}(\mathbb{R}^n)$上的作用定义为

$$\langle f,\varphi\rangle=\int_{\mathbb{R}^n}f(x)\varphi(x)\mathrm{d}x,\quad\forall\varphi\in\mathbb{S}(\mathbb{R}^n),$$

不难证明，$\langle f,\varphi\rangle$是$\mathbb{S}(\mathbb{R}^n)$上的连续线性泛函（证明留作习题），亦即任一个$f\in L^1(\mathbb{R}^n)$是一个 Schwartz 分布.

例 5.1.4 Heaviside 函数 $H(x)=\begin{cases}1, & x>0,\\ 0, & x<0.\end{cases}$

显然，$\langle H,\varphi\rangle=\int_{[0,+\infty)}\varphi(x)\mathrm{d}x$ 是$\mathbb{S}(\mathbb{R})$上的连续线性泛函，于是 $H(x)\in\mathbb{S}^*(\mathbb{R})$.

例 5.1.5 迪拉克(Dirac)分布 δ.

迪拉克分布 δ，作为$\mathbb{S}(\mathbb{R})$上的连续线性泛函 $\delta\in\mathbb{S}^*(\mathbb{R})$，定义为

$$\langle\delta,\varphi\rangle=\varphi(0),\quad\forall\varphi\in\mathbb{S}(\mathbb{R}).$$

δ **的线性**：$\forall\varphi_1,\varphi_2\in\mathbb{S}(\mathbb{R}),\alpha,\beta\in\mathbb{F}$，有

$$\langle\delta,\alpha\varphi_1+\beta\varphi_2\rangle=(\alpha\varphi_1+\beta\varphi_2)(0)=\alpha\varphi_1(0)+\beta\varphi_2(0)=\alpha\langle\delta,\varphi_1\rangle+\beta\langle\delta,\varphi\rangle;$$

δ **的连续性**：$\forall\{\varphi_n\}\subset\mathbb{S}(\mathbb{R}),\varphi\in\mathbb{S}(\mathbb{R})$，

$$\varphi_n\xrightarrow{\mathbb{S}(\mathbb{R})}\varphi\Rightarrow|\langle\delta,\varphi_n\rangle-\langle\delta,\varphi\rangle|=|\varphi_n(0)-\varphi(0)|\to0.$$

于是，我们得到了不能用积分形式表示的连续线性泛函.

δ **的支集**：$\mathrm{supp}\delta=\{0\}$.

由定义$\langle\delta,\varphi\rangle=\varphi(0)$，即得 $\mathrm{supp}\delta=\{0\}$.

例 5.1.6 迪拉克分布 δ 的导数.

由分布导数的定义，有

$$\langle\delta',\varphi\rangle=-\langle\delta,\varphi'\rangle=-\varphi'(0),\quad\forall\varphi\in\mathbb{S}(\mathbb{R}).$$

作为分布的 δ，它的导数也是一个分布，而且对每个 $\varphi\in\mathbb{S}(\mathbb{R})$作用的结果，就是$-\varphi'(0)$.

狄拉克分布 $\delta\in\mathbb{S}^*(\mathbb{R})$的 $p\in\mathbb{N}$ 阶导数为

$$\langle D^p\delta,\varphi\rangle=(-1)^p\varphi^{(p)}(0),\quad\forall\varphi\in\mathbb{S}(\mathbb{R}).$$

例 5.1.7 Heaviside 函数的导数.

由分布导数的定义，有

$$\langle H',\varphi\rangle=-\langle H,\varphi'\rangle=-\int_0^{+\infty}\varphi'(x)\mathrm{d}x=\varphi(0)=\langle\delta,\varphi\rangle,\quad\forall\varphi\in\mathbb{S}(\mathbb{R}),$$

从而 $H'=\delta$.

5. Schwartz 分布的分类

定义 5.1.6(正则分布、奇异分布) Schwartz 空间$\mathbb{S}(\mathbb{R}^n)$上的连续线性泛函 $T\in\mathbb{S}^*(\mathbb{R}^n)$可分为两类：

(1) 若分布 $T\in\mathbb{S}^*(\mathbb{R}^n)$对 $\varphi\in\mathbb{S}(\mathbb{R}^n)$的作用可表示为$\mathbb{R}^n$ 上的局部可积函数(函数在$\mathbb{R}^n$ 中的任意有限方体上都是 Lebesgue 可积的)$f\in L_{\mathrm{loc}}(\mathbb{R}^n)$的积分

$$\langle T,\varphi\rangle=\int_{\mathbb{R}^n}f(x)\varphi(x)\mathrm{d}x,\quad\forall\varphi\in\mathbb{S}(\mathbb{R}^n),$$

则称 T 为**正则分布(regular distribution)**；

(2) 若分布 $T\in\mathbb{S}^*(\mathbb{R}^n)$不是正则的，便称其为**奇异分布(singular distribution)**.

例如，因$\mathbb{S}(\mathbb{R}^n)\subset L^1(\mathbb{R}^n)\subset S^*(\mathbb{R}^n)$，所以 $f\in\mathbb{S}(\mathbb{R}^n)$与 $f\in L^1(\mathbb{R}^n)$时，f 都是正则分布. 但狄拉克分布 δ 是奇异分布.

有了 Schwartz 分布的分类，我们来说明前面第 2 小节中定义 Schwartz 分布的导数运算的思路与合理性.

为什么对于分布能定义导数呢？事实上，由于$\mathbb{S}(\mathbb{R})\subset\mathbb{S}^*(\mathbb{R})$，取 $f\in\mathbb{S}(\mathbb{R})$，则 $f'\in\mathbb{S}(\mathbb{R})$，并且 f 与 f'都是正则 Schwartz 分布. 于是，$\forall\varphi\in\mathbb{S}(\mathbb{R})$，由分部积分公式

$$\langle f',\varphi\rangle=\int_{-\infty}^{+\infty}f'(x)\varphi(x)\mathrm{d}x=f(x)\varphi(x)\Big|_{-\infty}^{+\infty}-\int_{-\infty}^{+\infty}f(x)\varphi'(x)\mathrm{d}x=-\langle f,\varphi'\rangle$$

得到$\langle f',\varphi\rangle=-\langle f,\varphi'\rangle$. 因此，当一个函数 f 不存在导数，或一个分布根本谈不上导数时，就可以用通常导数所应满足的表示式来定义它们的导数. 这就是用表示式

$$\langle\partial_{x_1}T,\varphi\rangle=-\langle T,\partial_{x_1}\varphi\rangle,\quad\forall\varphi\in\mathbb{S}(\mathbb{R}^n)$$

来定义分布导数的理由. 分布的其余运算请读者进行同样的推理与理解.

5.1.3 空间 $E(\mathbb{R}^n)$、$D(\mathbb{R}^n)$及其分布空间

5.1.1 节与 5.1.2 节讨论了 Schwartz 空间$\mathbb{S}(\mathbb{R}^n)$及其分布空间$\mathbb{S}^*(\mathbb{R}^n)$的性质. 实际上，常用的函数空间及其分布空间还有许多，本小节再介绍两个常用的空间及其分布空间.

1. 空间 $E(\Omega)\equiv C^\infty(\Omega)$及其分布空间

设 $\Omega\subseteq\mathbb{R}^n$ 是$\mathbb{R}^n$ 中的开集，$O(0,0,\cdots,0)\in\Omega$ 为原点，又设

$$E(\Omega)\equiv C^\infty(\Omega)=\{\partial^\beta\varphi\in C(\Omega):\forall\beta\in\mathbb{N}^n\}$$

是 $\Omega\subset\mathbb{R}^n$ 上任意次连续可微的函数类. 通常考虑集合 $E(\mathbb{R}^n)\equiv C^\infty(\mathbb{R}^n)$. 赋予 $E(\Omega)$自然拓扑如下：

取任意紧致集序列 $K_j\subset\Omega\subseteq\mathbb{R}^n$，$K_j\subset\mathring{K}_{j+1}$，使得 $\bigcup\limits_{j=1}^{+\infty}K_j=\Omega$. 对于 $\varphi\in E(\Omega)$，定义半范数

$$r_{m,j}(\varphi)=\sup_{\substack{x\in K,\\|\beta|\leqslant m}}|\partial^\beta\varphi(x)|,\quad m\in\mathbb{N},j\in\mathbb{Z}^+,\beta\in\mathbb{N}^n,$$

则$\{r_{m,j}(\varphi)\}$，$m\in\mathbb{N}$，$j\in\mathbb{Z}^+$是 $E(\Omega)$上的半范数族，使得$(E(\Omega),+,\alpha\cdot,\{r_{m,j}\})$成为数域$\mathbb{F}$上的 T_2 型、完备的、自反的拓扑线性空间.

$E(\Omega)$的共轭空间 $E^*(\Omega)$赋予如下的 w^* 拓扑：
对于序列$\{f_k\}\subset E^*(\Omega)$，有

① $f_k\xrightarrow{w^*}0(k\to+\infty)\Leftrightarrow\langle f_k,\varphi\rangle\to0(k\to+\infty)$，$\forall\varphi\in E(\Omega)$；

② $f\in E^*(\Omega)\Leftrightarrow\exists c>0,m\geqslant0$，紧致集 $K\subset\Omega$，使得

$$|\langle f,\varphi\rangle|\leqslant c\sup_{\substack{x\in K,\\|\beta|\leqslant m}}|\partial^\beta\varphi(x)|,\quad\forall\varphi\in E(\Omega).$$

定理 5.1.4（$E^*(\mathbb{R}^n)$拓扑结构定理） $(E^*(\mathbb{R}^n),+,\alpha\cdot,\tau_{w^*})$中的分布都具有紧致支集；反之亦然，亦即

$$f \in E^*(\mathbb{R}^n) \Leftrightarrow \mathrm{supp} f = K \text{ 是紧致集.}$$

例 5.1.8 $f\in L_{\mathrm{loc}}(\mathbb{R}^n)$，且$\mathrm{supp} f$ 是紧致集，则 $f\in E^*(\mathbb{R}^n)$. 线性泛函由下式确定：

$$\langle f,\varphi\rangle = \int_{\mathbb{R}^n} f(x)\varphi(x)\mathrm{d}x, \quad \forall \varphi \in E(\mathbb{R}^n).$$

例 5.1.9 迪拉克(Dirac)分布 $\delta\in E^*(\mathbb{R})$.

2. 空间 $D(\Omega)\equiv C_C^\infty(\Omega)$及其分布空间

1) $D(\Omega)$上的归纳极限拓扑

设 $\Omega\subseteq\mathbb{R}^n$ 是$\mathbb{R}^n$ 中的开集，$O(0,0,\cdots,0)\in\Omega$ 为原点，又设

$$D(\Omega) \equiv C_C^\infty(\Omega) = \{\varphi \in C^\infty(\Omega) ; \mathrm{supp}\varphi = K \subset \Omega, K \text{ 是紧致集}\}$$

是 $\Omega\subseteq\mathbb{R}^n$ 上任意次连续可导的、具有紧致支集的函数类. 通常也考虑集合 $D(\mathbb{R}^n)\equiv C_C^\infty(\mathbb{R}^n)$. 赋予归纳极限拓扑 τ(这里不详细叙述归纳极限拓扑的定义，只给出一种等价条件)：

对任意序列$\{\varphi_k\}\subset D(\Omega)$，$\varphi_k \xrightarrow{D(\Omega)} 0(k\to+\infty)$，当且仅当

① $\exists$紧致集 $K\subset\Omega$，使得$\mathrm{supp}\varphi_k\subset K$，$k\in\mathbb{Z}^+$；

② $\forall\beta\in\mathbb{N}^n$，使得 $\partial^\beta\varphi_k(x)\to 0$ 在 K 上一致成立.

于是，$(D(\Omega),+,\alpha\cdot,\tau)$成为数域 $\mathbb{F}$ 上的 T_2 型、完备的、自反的拓扑线性空间.

2) $D(\Omega)$的共轭空间 $D^*(\Omega)$上的拓扑

赋予 w^* 拓扑如下：

对于序列$\{f_k\}\subset D^*(\Omega)$，有

① $f_k \xrightarrow{w^*} 0(k\to+\infty) \Leftrightarrow \langle f_k,\varphi\rangle\to 0(k\to+\infty)$，$\forall\varphi\in D(\Omega)$；

② $\forall\beta\in\mathbb{N}^n$，$\partial^\beta f_k \xrightarrow{w^*} 0(k\to+\infty) \Leftrightarrow \langle\partial^\beta f_k,\varphi\rangle\to 0(k\to+\infty)$，$\forall\varphi\in D(\Omega)$.

例 5.1.10 $f\in L_{\mathrm{loc}}(\mathbb{R}^n)$，则 $f\in D^*(\mathbb{R}^n)$.

线性泛函由

$$\langle f,\varphi\rangle = \int_{\mathbb{R}^n} f(x)\varphi(x)\mathrm{d}x, \quad \forall \varphi \in D(\mathbb{R}^n)$$

确定，由于 $\varphi\in D(\mathbb{R}^n)$，$\mathrm{supp}\varphi$ 是紧致集，故上述积分有意义，且由 $f\in L_{\mathrm{loc}}(\mathbb{R}^n)$知，上述积分定义 $D(\mathbb{R}^n)$上的一个连续线性泛函.

例 5.1.11 $\delta\in D^*(\mathbb{R}^n)$.

显然，$\langle\delta,\varphi\rangle=\varphi(0)$，$\forall\varphi\in D(\mathbb{R}^n)$. 由于

$$D(\mathbb{R}^n) \equiv C_C^\infty(\mathbb{R}^n) \subset \mathbb{S}(\mathbb{R}^n) \subset C^\infty(\mathbb{R}^n) \equiv E(\mathbb{R}^n),$$

故 $D^*(\mathbb{R}^n)\supset\mathbb{S}^*(\mathbb{R}^n)\supset E^*(\mathbb{R}^n)$. (此包含关系的证明留作习题)

进而，由于 δ 分布的支集为$\mathrm{supp}\delta=\{0\}$，故 $\delta\in E^*(\mathbb{R}^n)\subset D^*(\mathbb{R}^n)$.

5.2 $L^p(\mathbb{R})(1\leqslant p\leqslant 2)$上的 Fourier 变换

5.2.1 $L^1(\mathbb{R})$上的 Fourier 变换

1. 预备知识

Fourier 变换理论是建立在 Lebesgue 测度与 Lebesgue 积分的基础上，这里先给出几个重要的不等式.

命题 5.2.1 设 $f\in L^p(\mathbb{R}^n), g\in L^q(\mathbb{R}^n), 1\leqslant p,q\leqslant +\infty, \dfrac{1}{p}+\dfrac{1}{q}=1$，则成立 Hölder 不等式

$$\|fg\|_{L^1(\mathbb{R}^n)}\leqslant \|f\|_{L^p(\mathbb{R}^n)}\|g\|_{L^q(\mathbb{R}^n)};$$

当 $p=q=2$ 时，上式成为 Schwartz 不等式

$$\|fg\|_{L^1(\mathbb{R}^n)}\leqslant \|f\|_{L^2(\mathbb{R}^n)}\|g\|_{L^2(\mathbb{R}^n)}.$$

命题 5.2.2 设 $p\geqslant 1, f,g\in L^p(\mathbb{R}^n)$，则有 Minkowski 不等式

$$\|f+g\|_{L^p(\mathbb{R}^n)}\leqslant \|f\|_{L^p(\mathbb{R}^n)}+\|g\|_{L^p(\mathbb{R}^n)}.$$

以上三个基本不等式经常用到. 读者可以在实变函数教程中查到它们的证明([6],[17]).

命题 5.2.3 卷积 $f*g(x)=\int_{\mathbb{R}^n}f(t)g(x-t)\mathrm{d}t$ 有如下性质：

(1) 当 $1\leqslant p,q\leqslant +\infty, \dfrac{1}{p}+\dfrac{1}{q}=1$，即 p,q 互为共轭数时，

① $f\in L^p(\mathbb{R}^n), g\in L^q(\mathbb{R}^n), 1\leqslant p,q\leqslant +\infty \Rightarrow f*g\in C(\mathbb{R}^n)$，亦即，$f*g(x)$是$\mathbb{R}^n$ 上的有界连续函数，且 $\|f*g\|_{C(\mathbb{R}^n)}\leqslant \|f\|_{L^p(\mathbb{R}^n)}\|g\|_{L^q(\mathbb{R}^n)}$；

② $f\in L^p(\mathbb{R}^n), g\in L^q(\mathbb{R}^n), 1< p,q< +\infty \Rightarrow f*g\in C_0(\mathbb{R}^n)$，这里 $C_0(\mathbb{R}^n)=\{f\in C(\mathbb{R}^n): \lim\limits_{|x|\to+\infty}f(x)=0\}$，且 $\|f*g\|_{C(\mathbb{R}^n)}\leqslant \|f\|_{L^p(\mathbb{R}^n)}\|g\|_{L^q(\mathbb{R}^n)}$；

(2) $f\in L^1(\mathbb{R}^n), g\in C_0(\mathbb{R}^n) \Rightarrow f*g\in C_0(\mathbb{R}^n)$，且 $\|f*g\|_{C(\mathbb{R}^n)}\leqslant \|f\|_{L^1(\mathbb{R}^n)}\|g\|_{C(\mathbb{R}^n)}$；

(3) $f\in L^p(\mathbb{R}^n), 1\leqslant p\leqslant +\infty, g\in L^1(\mathbb{R}^n) \Rightarrow f*g\in L^p(\mathbb{R}^n)$，且

$$\|f*g\|_{L^p(\mathbb{R}^n)}\leqslant \|f\|_{L^p(\mathbb{R}^n)}\|g\|_{L^1(\mathbb{R}^n)}.$$

2. Fourier 变换的定义

定义 5.2.1 函数 $f\in L^1(\mathbb{R}^n)$的 Fourier 变换定义为

$$f^?(\xi)=\int_{\mathbb{R}^n}f(x)\mathrm{e}^{-\mathrm{i}x\cdot\xi}\mathrm{d}x,\quad \xi\in\mathbb{R}^n,$$

其中 $x\cdot\xi=\sum\limits_{j=1}^{n}x_j\xi_j$ 是 $x=(x_1,x_2,\cdots,x_n)\in\mathbb{R}^n$ 与 $\xi=(\xi_1,\xi_2,\cdots,\xi_n)\in\mathbb{R}^n$ 的内积.

注 在一些 Fourier 分析的专著中,将 Fourier 变换定义为 $f^?(\xi)=\dfrac{1}{(\sqrt{2\pi})^n}\displaystyle\int_{\mathbb{R}^n} f(x)\mathrm{e}^{-\mathrm{i}x\cdot\xi}\mathrm{d}x,\xi\in\mathbb{R}^n$;那么,Fourier 逆变换(定义 5.2.2)就需定义为 $g^?(\xi)=\dfrac{1}{(\sqrt{2\pi})^n}\displaystyle\int_{\mathbb{R}^n} g(\boldsymbol{x})\mathrm{e}^{\mathrm{i}x\cdot\xi}\mathrm{d}\boldsymbol{x},\xi\in\mathbb{R}^n$,这样定义是为了某些结果的对称性,两种定义没有本质不同,只是在底空间 $\mathbb{R}^n$ 中选取了相差一个常数因子的测度而已. 但是,两种定义得到的性质大多数不变.

3. Fourier 变换的性质

以 $L^1(\mathbb{R})\equiv L^1(\mathbb{R}^1)$ 为例,叙述与证明性质与定理,但以下性质与定理对于 $L^1(\mathbb{R}^n)$,一般的 $n(n\geqslant 1)$,都成立.

1) 基本运算性质

定理 5.2.1 函数 $f\in L^1(\mathbb{R})$ 的 Fourier 变换的基本运算性质如下:

(1) $[f(\circ-h)]^?(\xi)=\mathrm{e}^{-ih\xi}f^?(\xi)$, $\xi\in\mathbb{R}$, $h\in\mathbb{R}$; (平移的 FT)

(2) $\tau_h f^?(\xi)=[\mathrm{e}^{ih\circ}f(\circ)]^?(\xi)$, $\xi\in\mathbb{R}$, $h\in\mathbb{R}$; (FT 的平移)

(3) $[\rho f(\rho\circ)]^?(\xi)=f^?\left(\dfrac{\xi}{\rho}\right)$, $\xi\in\mathbb{R}$, $\rho>0$; (伸缩的 FT)

(4) $\left[\overline{f(-\circ)}\right]^?(\xi)=\overline{f^?}(\xi)$, $\xi\in\mathbb{R}$. (反射的 FT)

根据定义就可以证明以上运算公式,请读者自行证明并熟练运用.

平移公式(1)也写为 $(\tau_h f)^?(\xi)=\mathrm{e}^{-ih\xi}f^?(\xi)$.

2) 分析性质

定理 5.2.2 (1) 函数 $f\in L^1(\mathbb{R})$ 的 Fourier 变换 $f^?(\xi)$ 是 $\mathbb{R}$ 上的一致连续函数,且 $f^?\in C(\mathbb{R})$;(2) Fourier 变换是 $L^1(\mathbb{R})$ 到 $L^\infty(\mathbb{R})$ 的一对一的连续线性算子;Fourier 变换视为算子?:$L^1(\mathbb{R})\to L^\infty(\mathbb{R})$,满足范数非增不等式 $\|f^?\|_{L^\infty(\mathbb{R})}\leqslant\|f\|_{L^1(\mathbb{R})}$,这里 $L^\infty(\mathbb{R})$ 是 $\mathbb{R}$ 上的本性有界函数空间,表示为

$$L^\infty(\mathbb{R})=\{f:\mathbb{R}\to\mathbb{R},\|f\|_{L^\infty(\mathbb{R})}<+\infty\},$$

其中 $\|f\|_{L^\infty(\mathbb{R})}=\operatorname{esssup}\{|f(x)|:x\in\mathbb{R}\}$ 称为 f 的本性上确界,就是 f 在空间 $L^\infty(\mathbb{R})$ 中的范数.

证 对于 $f\in L^1(\mathbb{R})$,其 Fourier 变换满足

$$|f^?(\xi+h)-f^?(\xi)|\leqslant\int_{\mathbb{R}}|\mathrm{e}^{-\mathrm{i}xh}-1||f(x)|\mathrm{d}x.$$

由 $f\in L^1(\mathbb{R})$,上式右边满足

$$\int_{\mathbb{R}}|\mathrm{e}^{-\mathrm{i}xh}-1||f(x)|\mathrm{d}x\leqslant M\int_{\mathbb{R}}|f(x)|\mathrm{d}x\leqslant M\|f\|_{L^1(\mathbb{R})},$$

故由 Lebesgue 控制收敛定理,立即得到

$$\lim_{|h|\to 0}|f^\wedge(\xi+h)-f^\wedge(\xi)|\leqslant\lim_{|h|\to 0}\int_{\mathbb{R}}|e^{-ihx}-1||f(x)|dx$$
$$=\int_{\mathbb{R}}\lim_{|h|\to 0}|e^{-ihx}-1||f(x)|dx=0$$

对于 $\xi\in\mathbb{R}$ 一致成立，所以 $f^\wedge(\xi)$是$\mathbb{R}$ 上的一致连续函数，且 $f^\wedge\in C(\mathbb{R})$.

进而，由 Fourier 变换的定义得到

$$|f^\wedge(\xi)|\leqslant\int_{\mathbb{R}}|f(x)e^{-ix\xi}|dx\leqslant\int_{\mathbb{R}}|f(x)|dx=\|f\|_{L^1(\mathbb{R})},$$

求本性上确界便得到 $\|f^\wedge\|_{L^\infty(\mathbb{R})}\leqslant\|f\|_{L^1(\mathbb{R})}$，此即算子的有界性(连续性).

3）Riemann-Lebesgue 引理

定理 5.2.3 函数 $f\in L^1(\mathbb{R})$的 Fourier 变换成立 Reimann-Lebesgue 引理

$$\lim_{|\xi|\to+\infty}f^\wedge(\xi)=0.$$

证 对于 $f\in L^1(\mathbb{R})$，由

$$f^\wedge(\xi)=\frac{1}{2}\int_{\mathbb{R}}\left[f(x)-f\left(x+\frac{\pi}{\xi}\right)\right]e^{-ix\xi}dx,$$

于是，利用 $f\in L^1(\mathbb{R})$的平均连续性，即 $\lim\limits_{h\to 0}\int_{\mathbb{R}}|f(x+h)-f(x)|dx=0$，得 $\lim\limits_{|\xi|\to+\infty}f^\wedge(\xi)=0$.

4）卷积的 Fourier 变换定理

定理 5.2.4 函数 $f,g\in L^1(\mathbb{R})$的卷积

$$f*g(x)=\int_{\mathbb{R}}f(t)g(x-t)dt$$

满足 $f*g\in L^1(\mathbb{R})$，且有 Fourier 变换公式

$$(f*g)^\wedge(\xi)=f^\wedge(\xi)g^\wedge(\xi),\quad \xi\in\mathbb{R}.$$

证 当 $f,g\in L^1(\mathbb{R})$时，有

$$\int_{\mathbb{R}}|f(t)g(x-t)|dx=|f(t)|\int_{\mathbb{R}}|g(x-t)|dx=|f(t)|\ \|g\|_{L^1(\mathbb{R})},$$

所以

$$\int_{\mathbb{R}}|f(t)g(x-t)|dx=|f(t)|\ \|g\|_{L^1(\mathbb{R})}\in L^1(\mathbb{R}).$$

于是，由 Fubini 定理，知

$$\int_{\mathbb{R}}|(f*g)(x)|dx\leqslant\int_{\mathbb{R}}\left|\int_{\mathbb{R}}|f(t)||g(x-t)|dt\right|dx$$
$$=\int_{\mathbb{R}}|f(t)|dt\int_{\mathbb{R}}|g(x-t)|dx=\|f\|_{L^1(\mathbb{R})}\|g\|_{L^1(\mathbb{R})}.$$

故 $f*g\in L^1(\mathbb{R})$. 进而，有

$$(f*g)^{?}(\xi)=\int_{\mathbb{R}}\left\{\int_{\mathbb{R}}f(t)g(x-t)\mathrm{d}t\right\}\mathrm{e}^{-\mathrm{i}x\xi}\mathrm{d}x$$

$$=\left\{\int_{\mathbb{R}}f(t)\mathrm{e}^{-\mathrm{i}t\xi}\mathrm{d}t\right\}\left\{\int_{\mathbb{R}}g\ (x-t)\mathrm{e}^{-\mathrm{i}(x-t)\xi}\mathrm{d}x\right\}=f^{?}(\xi)g^{?}(\xi).$$

定理得证.

此定理的数学意义在于 Fourier 变换将卷积运算(积分)转换为按点乘积(乘法)运算.

4. 导数的 Fourier 变换公式

"局部绝对连续函数类"$AC_{\mathrm{loc}}(\mathbb{R})$与"局部绝对 $k-1$ 阶连续函数类"$AC_{\mathrm{loc}}^{k-1}(\mathbb{R})$分别定义为

$$AC_{\mathrm{loc}}(\mathbb{R})=\left\{f:\forall x\in\mathbb{R},\exists g\in L_{\mathrm{loc}}(\mathbb{R}),\text{s. t. }f(x)=f(0)+\int_0^x g(u)\mathrm{d}u\right\},$$

$$AC_{\mathrm{loc}}^{k-1}(\mathbb{R})=\{f:\forall x\in\mathbb{R},\text{有 }f(x)=a_0+F(x)\},$$

其中 $F(x)$满足：$\exists g\in L_{\mathrm{loc}}(\mathbb{R})$，$\exists$ 常数 $a_0,a_1,\cdots,a_{k-1}$，使得

$$F(x)=\int_0^x\mathrm{d}u_1\left[a_1+\int_0^{u_1}\mathrm{d}u_2+\left[a_2+\cdots+\int_0^{u_{k-2}}\mathrm{d}u_{k-2}\left(a_{k-1}+\int_0^{u_{k-1}}g(u_k)\mathrm{d}u_k\right)\cdots\right]\right]\mathrm{d}u.$$

定理 5.2.5　函数 $f\in L^1(\mathbb{R})\cap AC_{\mathrm{loc}}(\mathbb{R})$，若导数 $f'\in L^1(\mathbb{R})$，则其 Fourier 变换公式为

$$(f')^{?}(\xi)=\mathrm{i}\xi f^{?}(\xi),\quad \xi\neq 0;$$

进而，若 $f\in L^1(\mathbb{R})\cap AC_{\mathrm{loc}}^{k-1}(\mathbb{R})$，$f^{(k)}\in L(\mathbb{R})$，则

$$(f^{(k)})^{?}(\xi)=(\mathrm{i}\xi)^k f^{?}(\xi),\quad \xi\neq 0.$$

证明可参看[4].

5. Fourier 变换的导数公式

定理 5.2.6　函数 $f\in L^1(\mathbb{R})$，$x^k f\in L^1(\mathbb{R})$，则 f 的 Fourier 变换 $f^{?}$ 具有属于 $C_0(\mathbb{R})$ 的 k 阶导数，且有公式

$$(f^{?})^{(r)}(\xi)=(-i)^r\int_{\mathbb{R}}x^r f\ (x)\mathrm{e}^{-\mathrm{i}\xi x}\mathrm{d}x,\quad r=1,2,\cdots,k,\ \xi\in\mathbb{R}.$$

6. Fourier 变换的 Parseval 公式

定理 5.2.7　(1) 若 $f,g\in L^1(\mathbb{R})$，则成立 Fourier 变换的 Parseval 公式

$$\int_{\mathbb{R}}f(t)g^{?}(t)\mathrm{d}t=\int_{\mathbb{R}}f^{?}(t)g(t)\mathrm{d}t;$$

(2) 若 $f,g\in L^1(\mathbb{R})$，$g^{?}\in L^1(\mathbb{R})$，则成立 Fourier 变换的 Plancherel 型公式

$$\int_{\mathbb{R}}f(x)\ \overline{g(x)}\mathrm{d}x=\frac{1}{2\pi}\int_{\mathbb{R}}f^{?}(\xi)\ \overline{g^{?}(\xi)}\mathrm{d}\xi,$$

并且

$$f*g(x)=\frac{1}{2\pi}\int_{\mathbb{R}}f^{?}(\xi)g^{?}(\xi)\mathrm{e}^{\mathrm{i}x\xi}\mathrm{d}\xi,\quad \text{a. e. } x\in\mathbb{R}.$$

7. Fourier 逆变换的定义及性质

定义 5.2.2(逆变换) 函数 $g\in L^1(\mathbb{R}^n)$的 Fourier 逆变换定义为

$$g^{?}(x)=\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}g(\xi)\mathrm{e}^{\mathrm{i}x\cdot\xi}\mathrm{d}\xi,\quad x\in\mathbb{R}^n,$$

其中 $x\cdot\xi=\sum_{j=1}^{n}x_j\xi_j$ 是 $x=(x_1,x_2,\cdots,x_n)\in\mathbb{R}^n$ 与 $\xi=(\xi_1,\xi_2,\cdots,\xi_n)\in\mathbb{R}^n$ 的内积. 当 $n=1$ 时,逆变换呈 $g^{?}(x)=\frac{1}{2\pi}\int_{\mathbb{R}}g(\xi)\mathrm{e}^{\mathrm{i}x\xi}\mathrm{d}\xi,x\in\mathbb{R}$ 形式.

为叙述简单,以下仍对 $L^1(\mathbb{R})$情况叙述有关定理.

定理 5.2.8(反演公式) (1) 若函数 $f,f^{?}\in L^1(\mathbb{R})$,则

$$f(x)=\frac{1}{2\pi}\int_{\mathbb{R}}f^{?}(\xi)\mathrm{e}^{\mathrm{i}\xi x}\mathrm{d}\xi,\quad \text{a. e. } x\in\mathbb{R};$$

(2) 若函数 $f\in L^1(\mathbb{R})\cap C(\mathbb{R}),f^{?}\in L^1(\mathbb{R})$,则

$$f(x)=\frac{1}{2\pi}\int_{\mathbb{R}}f^{?}(\xi)\mathrm{e}^{\mathrm{i}\xi x}\mathrm{d}\xi,\quad \forall x\in\mathbb{R};$$

(3) 若函数 $f\in L^1(\mathbb{R}),f^{?}(\xi)=0$,则

$$f(x)=0,\quad \text{a. e. } x\in\mathbb{R};$$

(4) 若函数 $f\in L^1(\mathbb{R})$,则引进“求和因子”后,就可求得 $f(x)$,a. e. $x\in\mathbb{R}$,

① Cesaro 求和因子$\left(1-\frac{|\xi|}{\rho}\right)$:

$$\lim_{\rho\to 0}\frac{1}{2\pi}\int_{-\rho}^{\rho}\left(1-\frac{|\xi|}{\rho}\right)f^{?}(\xi)\mathrm{e}^{\mathrm{i}\xi x}\mathrm{d}\xi=f(x),\quad \text{a. e. } x\in\mathbb{R};$$

② Abel 求和因子 $\exp\left\{-\frac{|\xi|}{\rho}\right\}$:

$$\lim_{\rho\to+\infty}\frac{1}{2\pi}\int_{\mathbb{R}}\exp\left\{-\frac{|\xi|}{\rho}\right\}f^{?}(\xi)\mathrm{e}^{\mathrm{i}\xi x}\mathrm{d}\xi=f(x),\quad \text{a. e. } x\in\mathbb{R};$$

③ Gauss 求和因子 $\exp\left\{-\left(\frac{\xi}{\rho}\right)^2\right\}$:

$$\lim_{\rho\to+\infty}\frac{1}{2\pi}\int_{\mathbb{R}}\exp\left\{-\left(\frac{\xi}{\rho}\right)^2\right\}f^{?}(\xi)\mathrm{e}^{\mathrm{i}\xi x}\mathrm{d}\xi=f(x),\quad \text{a. e. } x\in\mathbb{R}.$$

证明可参看[4].

5.2.2 $L^2(\mathbb{R})$上的 Fourier 变换

大家知道,$L^2([a,b])\subset L^1([a,b])$,但是 $L^2(\mathbb{R})$与 $L^1(\mathbb{R})$的关系却并非如此,例如

函数

$$f(x)=\frac{1}{\sqrt{x}(1+|\ln|x||)}\in L^2(\mathbb{R}),$$

却有

$$f(x)=\frac{1}{\sqrt{x}(1+|\ln|x||)}\notin L^p(\mathbb{R}),\quad p\neq 2.$$

因此 Fourier 系数(有限 Fourier 变换)的定义只需对 $f\in L^1([a,b])$而作,但对于 Fourier 变换,在 $L^2(\mathbb{R})$情况,就必须"另起炉灶".

1. 两个预备命题

命题 5.2.4 设 $f\in L^1(\mathbb{R})\cap L^2(\mathbb{R})$,则 $f^?(\xi)\in L^2(\mathbb{R})$,且 $\|f^?\|_{L^2(\mathbb{R})}=\sqrt{2\pi}\,\|f\|_{L^2(\mathbb{R})}$.

证 令 $g(x)=\overline{f(-x)}$,$h(x)=(f*g)(x)$,则 $g(x)\in L^1(\mathbb{R})\cap L^2(\mathbb{R})$. 由命题 5.2.3 卷积性质(1)之② 与(3)得到

$$h=f*g\in L^1(\mathbb{R})\cap L^2(\mathbb{R}),$$

故 h 的 Fourier 变换有意义,且由定理 5.2.8 的反演公式(4)之②,并令 $x=0$,得

$$\lim_{\rho\to+\infty}\frac{1}{2\pi}\int_{\mathbb{R}}\exp\left\{-\frac{|\xi|}{\rho}\right\}h^?(\xi)\mathrm{d}\xi=h(0).\tag{5.2.1}$$

再据卷积 Fourier 变换公式 $(f*g)^?(\xi)=f^?(\xi)g^?(\xi)$与反射 Fourier 变换公式

$$g^?(\xi)=\left[\overline{f(-\circ)}\right]^?(\xi)=\overline{f^?(\xi)},$$

得

$$h^?(\xi)=f^?(\xi)g^?(\xi)=f^?(\xi)\,\overline{f^?(\xi)}=|f^?(\xi)|^2,\tag{5.2.2}$$

于是,(5.2.1)式中的被积函数为正函数,而且当 $\rho\to+\infty$时,$\exp\left\{-\frac{|\xi|}{\rho}\right\}h^?(\xi)$单调递减趋向于 $h^?(\xi)$,应用 Lebesgue 单调收敛定理,得到表示式 $h(0)=\frac{1}{2\pi}\int_{\mathbb{R}}\lim_{\rho\to+\infty}\exp\left\{-\frac{|\xi|}{\rho}\right\}h^?(\xi)\mathrm{d}\xi=\frac{1}{2\pi}\int_{\mathbb{R}}h^?(\xi)\mathrm{d}\xi$,亦即

$$\int_{\mathbb{R}}h^?(\xi)\mathrm{d}\xi=2\pi h(0),$$

所以 $h^?\in L^1(\mathbb{R})$. 另一方面,有

$$\begin{aligned}h(0)=(f*g)(0)&=\int_{\mathbb{R}}f(t)g(0-t)\mathrm{d}t\\&=\int_{\mathbb{R}}f(t)\,\overline{f(t)}\mathrm{d}t=\int_{\mathbb{R}}|f(t)|^2\mathrm{d}t=\|f\|_{L^2(\mathbb{R})}^2,\end{aligned}$$

结合(5.2.2)式得到 $\|f^?\|_{L^2(\mathbb{R})}^2=\int_{\mathbb{R}}|f^?(\xi)|^2\mathrm{d}\xi=\int_{\mathbb{R}}h^?(\xi)\mathrm{d}\xi=2\pi h(0)$,最后得 $\|f^?\|_{L^2(\mathbb{R})}=$

$\sqrt{2\pi}\,\|f\|_{L^2(\mathbb{R})}$.

命题 5.2.5 设 $f\in L^2(\mathbb{R})$，令 $f_\rho(x)=\begin{cases}f(x), & |x|\leqslant\rho,\\ 0, & |x|>\rho\end{cases}$ $(\rho>0)$，则 $f_\rho\in L^1(\mathbb{R})\cap L^2(\mathbb{R})$，且 $\forall\rho>0$，有 $f^{?}_\rho\in L^2(\mathbb{R})$；进而，存在惟一的 $g\in L^2(\mathbb{R})$，使得

$$\lim_{\rho\to+\infty}\|f^{?}_\rho-g\|_{L^2(\mathbb{R})}=0.$$

证 由 Hölder 不等式得到，对于所有 $\rho>0$，有 $f_\rho\in L^1(\mathbb{R})\cap L^2(\mathbb{R})$. 再由命题 5.2.4，得 $f^{?}_\rho\in L^2(\mathbb{R})$，并且 $f^{?}_{\rho_1}-f^{?}_{\rho_2}$ 是 $f_{\rho_1}-f_{\rho_2}$ 的 Fourier 变换；当 $\rho_1<\rho_2$ 时，有

$$\|f^{?}_{\rho_1}-f^{?}_{\rho_2}\|^2_{L^2(\mathbb{R})}=(2\pi)\|f_{\rho_1}-f_{\rho_2}\|^2_{L^2(\mathbb{R})}=(2\pi)\int_{(-\rho_2,-\rho_1)}|f(x)|^2\mathrm{d}x+(2\pi)\int_{(\rho_1,\rho_2)}|f(x)|^2\mathrm{d}x.$$

由于 $f\in L^2(\mathbb{R})$，上式右边的两个积分都趋于 0，于是，$\lim\limits_{\rho_1,\rho_2\to+\infty}\|f^{?}_{\rho_1}-f^{?}_{\rho_2}\|_{L^2(\mathbb{R})}=0$，从而，对于 $\rho\to+\infty$ 的任意子序列 $\{k_n\}(k_n\in\mathbb{Z}^+)$，Fourier 变换 $f^{?}_{k_n}\in L^2(\mathbb{R})$ 是 $L^2(\mathbb{R})$ 中的 Cauchy 序列；据 $L^2(\mathbb{R})$ 的完备性，存在惟一的函数 $g\in L^2(\mathbb{R})$，使得 $\lim\limits_{\rho\to+\infty}\|f^{?}_\rho-g\|_{L^2(\mathbb{R})}=0$，命题得证.

2. $L^2(\mathbb{R}^n)$函数的 Fourier 变换的定义

定义 5.2.3($L^2(\mathbb{R}^n)$函数的 Fourier 变换) 函数 $f\in L^2(\mathbb{R}^n)$ 的 Fourier 变换 $F^2(f)$ 定义为

$$[F^2(f)](\xi)=\underset{\rho\to+\infty}{\overset{(2)}{\mathrm{l.i.m.}}}\,f^{?}_\rho(\xi)=\underset{\rho\to+\infty}{\overset{(2)}{\mathrm{l.i.m.}}}\int_{(-\rho,\rho)^n}f(x)\mathrm{e}^{-\mathrm{i}x\cdot\xi}\mathrm{d}x,\quad \xi\in\mathbb{R}^n,$$

其中 $x\cdot\xi=\sum\limits_{j=1}^n x_j\xi_j$ 是 $x=(x_1,x_2,\cdots,x_n)\in\mathbb{R}^n$ 与 $\xi=(\xi_1,\xi_2,\cdots,\xi_n)\in\mathbb{R}^n$ 的内积；$\overset{(2)}{\mathrm{l.i.m.}}$ 是"limit in mean in L^2-sense".

这样，$f\in L^2(\mathbb{R}^n)$ 的 Fourier 变换是由下式惟一确定的函数 $g\equiv[F^2(f)]\in L^2(\mathbb{R}^n)$：

$$\lim_{\rho\to+\infty}\left\|F^2(f)(\circ)-\int_{(-\rho,\rho)^n}f(x)\mathrm{e}^{-\mathrm{i}x\cdot\circ}\mathrm{d}x\right\|_{L^2(\mathbb{R}^n)}=0.$$

注意，n 维方体 $(-\rho,\rho)^n$ 换为 n 维长方体 $(-\rho_1,\rho_1)\times\cdots\times(-\rho_n,\rho_n)$ 时，定义等价.

3. $L^2(\mathbb{R})$函数的 Fourier 变换的性质

在不发生混淆的情况下，对于 $f\in L^2(\mathbb{R})$，仍记 $f^{?}(\xi)\equiv(F^2(f))(\xi)$. 关于 $L^2(\mathbb{R})$ 函数 Fourier 变换的运算性质，与定理 5.2.1 相同，并由后面的定理 5.2.10 表述. 其分析性质如下.

定理 5.2.9(Plancherel 定理) 函数 $f\in L^2(\mathbb{R})$ 的 Fourier 变换是 $L^2(\mathbb{R})$ 到 $L^2(\mathbb{R})$ 的一对一的连续线性算子；对于 $f\in L^2(\mathbb{R})$，存在惟一的 $L^2(\mathbb{R})$ 函数 $[F^2(f)](\xi)$，满足范数等式

$$\|F^2(f)\|_{L^2(\mathbb{R})}=\sqrt{2\pi}\,\|f\|_{L^2(\mathbb{R})}.$$

证 $F^2(f)$显然是线性算子 $F^2(\alpha f+\beta g)=\alpha F^2(f)+\beta F^2(g)$，$f,g\in L^2(\mathbb{R}^n)$，$\alpha,\beta\in\mathbb{R}$.

由 Fourier 变换的定义，$\lim\limits_{\rho\to+\infty}\|F^2(f)-f_\rho^?\|_{L^2(\mathbb{R})}=0$，此式蕴含

$$\lim_{\rho\to+\infty}\|f_\rho^?\|_{L^2(\mathbb{R})}=\|F^2(f)\|_{L^2(\mathbb{R})};\tag{5.2.3}$$

另一方面，由 f_ρ 的定义，有

$$\lim_{\rho\to+\infty}\|f_\rho\|_{L^2(\mathbb{R})}=\|f\|_{L^2(\mathbb{R})}.\tag{5.2.4}$$

再由 $f_\rho\in L^1(\mathbb{R})\cap L^2(\mathbb{R})$，则由命题 5.2.4 得到

$$\|f_\rho^?\|_{L^2(\mathbb{R})}=\sqrt{2\pi}\,\|f_\rho\|_{L^2(\mathbb{R})},\tag{5.2.5}$$

对(5.2.5)式两边取极限，利用(5.2.3)式、(5.2.4)式得到 $\|F^2(f)\|_{L^2(\mathbb{R})}=\sqrt{2\pi}\,\|f\|_{L^2(\mathbb{R})}$，此即算子的范数等式与有界性.

注 在 Fourier 变换与逆变换的定义由对称形式给出时，范数等式化为保范性等式 $\|F^2(f)\|_{L^2(\mathbb{R})}=\|f\|_{L^2(\mathbb{R})}$.

5.2.3 $L^p(\mathbb{R})(1<p<2)$上的 Fourier 变换

对于 $L^p(\mathbb{R})(1<p<2)$ 中函数的 Fourier 变换，要涉及更多的理论，例如要用到 Riesz-Thorin 凸性定理、算子的型等，我们不再详述其证明，只讲思路与结果，并仍以 $n=1$ 为例.

1. 两个预备命题

记 $S_{00}(\mathbb{R})$为$\mathbb{R}$上的简单函数类，亦即，S_{00}中的函数在$\mathbb{R}$中的有限个测度为有限的集合上取有限值(如图 5.2.1 所示)，

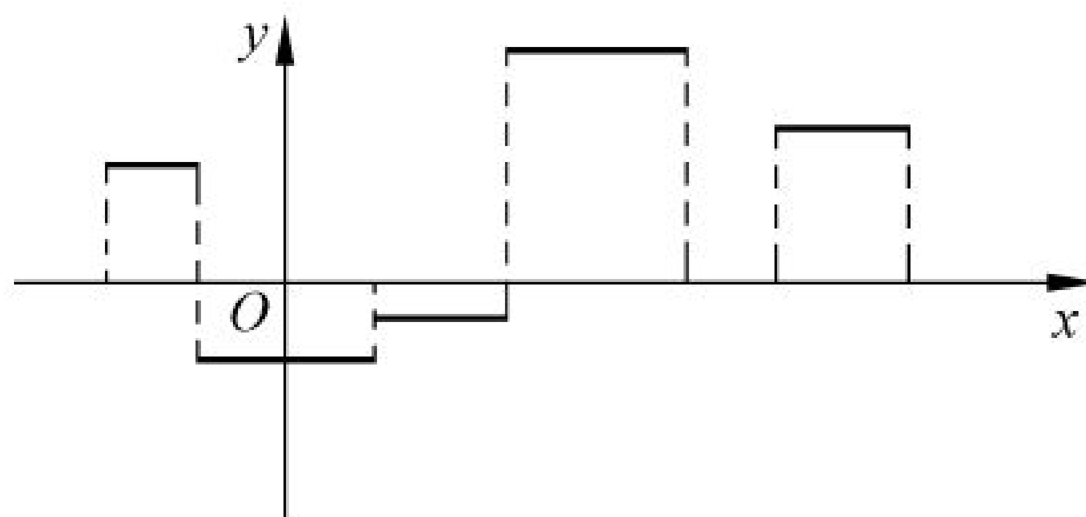


图 5.2.1 简单函数

$$h\in S_{00}(\mathbb{R})\Rightarrow h(x)=\sum_{j=1}^{m}c_j\chi_{e_j}(x),\quad x\in\mathbb{R},$$

其中 $-\infty<c_1<c_2<\cdots<c_m<+\infty$，$e_j=\{x\in\mathbb{R}:h(x)=c_j\}$，$0\leqslant m(e_j)<+\infty$，这里 $m(e_j)$是 e_j 的 Lebesgue 测度. 所以，集合 e_j 的个数、测度 $m(e_j)$、取值 c_j 都是有限的，$j=1,2,\cdots,m$.

命题 5.2.6($S_{00}(\mathbb{R})$**的性质**) (1) $S_{00}(\mathbb{R})\subset L^p(\mathbb{R})$，且 $S_{00}(\mathbb{R})$在 $L^p(\mathbb{R})(1\leqslant p<+\infty)$中稠密；亦即 $\forall\varepsilon>0$，$\forall f\in L^p(\mathbb{R})\Rightarrow\exists h\in S_{00}(\mathbb{R})$，s. t. $\|f-h\|_{L^p(\mathbb{R})}<\varepsilon$；

(2) 若$\mathrm{supp}f$ 为紧致集，则存在序列$\{h_n\}\subset S_{00}(\mathbb{R})$，s. t. $\lim\limits_{n\to+\infty}\|f-h_n\|_{L^p(\mathbb{R})}=0$，其中 $\mathrm{supp}h_n\subset\mathrm{supp}f$ 对每个 $n\in\mathbb{Z}^+$ 成立；

(3) 若 $h\in S_{00}(\mathbb{R})$，则对于 $1\leqslant p\leqslant 2$，有 $h^?\in L^q(\mathbb{R})$，$\dfrac{1}{p}+\dfrac{1}{q}=1$，且 $\|h^?\|_{L^q(\mathbb{R})}\leqslant\|h\|_{L^p(\mathbb{R})}$.

注 $h\in S_{00}(\mathbb{R})$在所有 $L^p(\mathbb{R})(1\leqslant p<+\infty)$中，所以它有以下几种"身份"：

① 作为 $h\in S_{00}(\mathbb{R})$，由(3)，$h^{?}\in L^q(\mathbb{R})$，$\dfrac{1}{p}+\dfrac{1}{q}=1$；并且，由(1)，$S_{00}(\mathbb{R})\subset L^p(\mathbb{R})$在 $L^p(\mathbb{R})$稠密，故以后将定义的 $f\in L^p(\mathbb{R})(1<p<2)$的 Fourier 变换(定义 5.2.4)应当属于 $L^q(\mathbb{R})\left(\dfrac{1}{p}+\dfrac{1}{q}=1\right)$，且 $2<q<+\infty$；

② 作为 $h\in L^1(\mathbb{R})$，$h^{?}$ 可按$L^1(\mathbb{R})$定义 $Fh(\xi)\equiv h^{?}(\xi)=\int_{\mathbb{R}}h(x)\mathrm{e}^{-\mathrm{i}x\xi}\mathrm{d}x$，并且 $h^{?}\in C_0(\mathbb{R})$ 满足 $\|h^{?}\|_{L^\infty(\mathbb{R})}\leqslant\|h\|_{L^1(\mathbb{R})}$，称 Fourier 变换算子?为$(1,\infty)$型算子；

③ 作为 $h\in L^2(\mathbb{R})$，则有 $\|h^{?}\|_{L^2(\mathbb{R})}=\sqrt{2\pi}\,\|h\|_{L^2(\mathbb{R})}$，称 Fourier 变换算子?为$(2,2)$型算子.

这样，线性算子?是 $S_{00}(\mathbb{R})$上的"$(1,\infty)$型"与"$(2,2)$型"的，据 R-T 凸性定理(命题 5.2.7)，Fourier 变换?是(p,q)型线性算子，亦即成立不等式 $\|h^{?}\|_{L^q(\mathbb{R})}\leqslant\|h\|_{L^p(\mathbb{R})}$.

下述 Riesz-Thorin 凸性定理，是调和分析中的重要定理，可以对很一般的空间成立. 我们仅对 $L^p(\mathbb{R})$情形叙述. 对一般情形，读者可参考有关文献.

命题 5.2.7(Riesz-Thorin 凸性定理) 设 $T:L^p(\mathbb{R})\to L^r(\mathbb{R})(1\leqslant p,r\leqslant+\infty)$是线性算子，并且是$(p_0,r_0)$型算子，即满足不等式

$$\|Tf\|_{L^{r_0}(\mathbb{R})}\leqslant M_0\,\|f\|_{L^{p_0}(\mathbb{R})},$$

其中 $1\leqslant p_0,r_0\leqslant+\infty$；同时也是$(p_1,r_1)$型算子，即满足不等式

$$\|Tf\|_{L^{r_1}(\mathbb{R})}\leqslant M_1\,\|f\|_{L^{p_1}(\mathbb{R})},$$

其中 $1\leqslant p_1,r_1\leqslant+\infty$. 则算子 T 是(p,r)型的，且满足不等式

$$\|Tf\|_{L^r(\mathbb{R})}\leqslant M\,\|f\|_{L^p(\mathbb{R})},$$

其中 $\dfrac{1}{p}=\dfrac{1-t}{p_0}+\dfrac{1}{p_1}$，$\dfrac{1}{r}=\dfrac{1-t}{r_0}+\dfrac{1}{r_1}$，对任意 $t\in[0,1]$成立，并且 $M\leqslant M_0^{1-t}M_1^t$.

图 5.2.2 中所画坐标轴是$\dfrac{1}{p}$与$\dfrac{1}{r}$，对角线(粗线)上的点$\left(\dfrac{1}{p},\dfrac{1}{p}\right)$对应于$(p,p)$型；反对角线(细线)上的点$\left(\dfrac{1}{p},\dfrac{1}{q}\right)$，$\left(\dfrac{1}{p}+\dfrac{1}{q}=1\right)$，对应于$(p,q)$型；阴影部分的点$\left(\dfrac{1}{p},\dfrac{1}{r}\right)$则对应于$(p,r)$型；当 $p\leqslant r$ 时，$\left(\dfrac{1}{p},\dfrac{1}{r}\right)$出现在阴影部分.

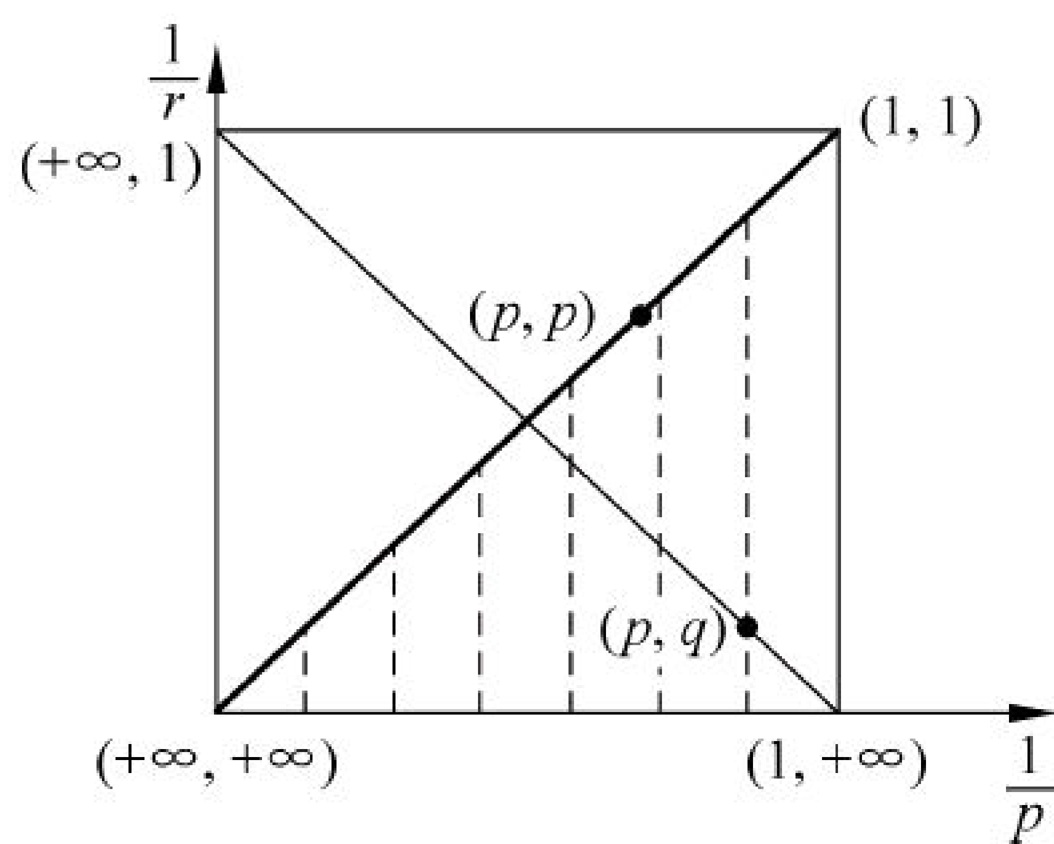


图 5.2.2 Riesz-Thorin 定理示意图

于是，对于 $1<p<2$，只能有 $2<q<+\infty$，这就是我们考虑的 $L^p(\mathbb{R}^n)(1<p<2)$中函数的 Fourier 变换，作为算子 $F^p:L^p(\mathbb{R}^n)\to L^q(\mathbb{R})(1<p<2)$，只能有 $2<q<+\infty$的原因.

2. $L^p(\mathbb{R}^n)(1<p<2)$函数的 Fourier 变换的定义

与 $L^2(\mathbb{R}^n)$情形类似,有如下定义.

定义 5.2.4 函数 $f\in L^p(\mathbb{R}^n)(1<p<2)$的 Fourier 变换 $F^p(f)$定义为

$$[F^p(f)](\xi)=\underset{\rho\to+\infty}{\overset{(q)}{\text{l.i.m.}}}\, f_\rho^?(\xi)=\underset{\rho\to+\infty}{\overset{(q)}{\text{l.i.m.}}}\int_{(-\rho,\rho)^n} f(x)\mathrm{e}^{-\mathrm{i}x\cdot\xi}\mathrm{d}x,\quad \xi\in\mathbb{R}^n,$$

其中 $x\cdot\xi=\sum_{j=1}^{n}x_j\xi_j$ 是 $x=(x_1,x_2,\cdots,x_n)\in\mathbb{R}^n$ 与 $\xi=(\xi_1,\xi_2,\cdots,\xi_n)\in\mathbb{R}^n$ 的内积,$\frac{1}{p}+\frac{1}{q}=1$,$\overset{(q)}{\text{l.i.m.}}$ 是“limit in mean in L^q-sense”.

这样,$f\in L^p(\mathbb{R}^n)$的 Fourier 变换是由 $g\equiv[F^p(f)](1<p<2)$惟一确定的函数 $g\in L^q(\mathbb{R}^n)$,满足

$$\lim_{\rho\to+\infty}\left\|F^p(f)(\circ)-\int_{(-\rho,\rho)^n}f(x)\mathrm{e}^{-\mathrm{i}x\cdot\circ}\mathrm{d}x\right\|_{L^q(\mathbb{R}^n)}=0.$$

注意,n 维方体$(-\rho,\rho)^n$ 换为 n 维长方体$(-\rho_1,\rho_1)\times\cdots\times(-\rho_n,\rho_n)$时,定义等价.

3. $L^p(\mathbb{R})(1<p<2)$函数的 Fourier 变换的性质

以下也均讨论 $n=1$ 情形. 在不发生混淆的情况下,对于 $f\in L^p(\mathbb{R})(1<p\leqslant2)$,仍记其 Fourier 变换为 $f^?(\xi)\equiv[F^p(f)](\xi)$. 利用命题 5.2.6 与命题 5.2.7,就可证明以下定理.

定理 5.2.10(运算性质) 设 $f\in L^p(\mathbb{R})(1\leqslant p\leqslant2)$,

(1) $(\tau_hf)^?(\xi)=\mathrm{e}^{-\mathrm{i}h\xi}f^?(\xi)$, $\xi\in\mathbb{R}$, $h\in\mathbb{R}$; (平移的 FT)

(2) $\tau_hf^?(\xi)=[\mathrm{e}^{\mathrm{i}h\circ}f(\circ)]^?(\xi)$, $\xi\in\mathbb{R}$, $h\in\mathbb{R}$; (FT 的平移)

(3) $[\rho f(\rho\circ)]^?(\xi)=f^?\left(\dfrac{\xi}{\rho}\right)$, $\xi\in\mathbb{R}$, $\rho>0$; (伸缩的 FT)

(4) $\left[\overline{f(-\circ)}\right]^?(\xi)=\overline{f^?(\xi)}$, $\xi\in\mathbb{R}$. (反射的 FT)

此定理在 $p=1$ 时,就是定理 5.2.1.

定理 5.2.11(分析性质) 函数 $f\in L^p(\mathbb{R})(1<p<2)$的 Fourier 变换是 $L^p(\mathbb{R})$到 $L^q(\mathbb{R})$的一对一的连续线性算子;对于 $f\in L^p(\mathbb{R})(1<p<2)$,存在惟一的 $L^q(\mathbb{R})(2<q<+\infty)$函数,记为$[F^p(f)]\in L^q(\mathbb{R})\left(\dfrac{1}{p}+\dfrac{1}{q}=1\right)$,成立范数非增的 Titchmarsh 不等式

$$\|F^p(f)\|_{L^q(\mathbb{R})}\leqslant\|f\|_{L^p(\mathbb{R})}.$$

因此,算子 F^p 是(p,q)型的(不等式也包括了 $1\leqslant p\leqslant2$ 情形). 并且,若 f 同时属于 $L^1(\mathbb{R})$与 $L^p(\mathbb{R})(1<p\leqslant2)$,即 $f\in L^1(\mathbb{R})\cap L^p(\mathbb{R})(1<p\leqslant2)$,则

$$f^?(\xi)\equiv(F^1f)(\xi)=(F^pf)(\xi),\quad \text{a.e.}\,\xi\in\mathbb{R}.$$

定理 5.2.12(卷积公式) 函数 $f\in L^p(\mathbb{R})(1<p\leqslant 2)$,$g\in L^1(\mathbb{R})$的卷积

$$f*g(x)=\int_{\mathbb{R}}f(t)g(x-t)\mathrm{d}t$$

满足 $f*g\in L^p(\mathbb{R})$,且

$$F^p(f*g)(\xi)=(F^pf)(\xi)\cdot(F^1g)(\xi),\quad \xi\in\mathbb{R},$$

简记为

$$(f*g)^{?}(\xi)=f^{?}(\xi)g^{?}(\xi),\quad \xi\in\mathbb{R}.$$

定理 5.2.13(导数的 FT) 函数 $f\in L^p(\mathbb{R})\cap AC_{\mathrm{loc}}^{r-1}(\mathbb{R})(1<p\leqslant 2)$且导数 $f^{(r)}\in L^p(\mathbb{R})$,$r\in\mathbb{Z}^+$,则导数的 Fourier 变换公式为

$$(f^{(r)})^{?}(\xi)=(\mathrm{i}\xi)^r f^{?}(\xi),\quad \text{a.e.}\,\xi\in\mathbb{R}.$$

定理 5.2.14(FT 的导数) 函数 $f\in L^p(\mathbb{R})(1<p\leqslant 2)$,且$(\mathrm{i}x)^r\in L^1(\mathbb{R})$,$r\in\mathbb{Z}^+$,则 $L^p(\mathbb{R})$上 Fourier 变换的 r 阶导数,作为 $L^q(\mathbb{R})$中的函数$(f^{?})^{(r)}(\xi)$,几乎处处存在,且有公式

$$(f^{?})^{(r)}(\xi)=(-\mathrm{i})^r(\circ^r f(\circ))^{?}(\xi),\quad \text{a.e.}\,\xi\in\mathbb{R}.$$

4. Fourier 变换的 Parseval 公式

定理 5.2.15 (1) 若 $f,g\in L^p(\mathbb{R})(1<p\leqslant 2)$,则成立 Fourier 变换的 Parseval 公式

$$\int_{\mathbb{R}}f(t)g^{?}(t)\mathrm{d}t=\int_{\mathbb{R}}f^{?}(t)g(t)\mathrm{d}t;$$

(2) 若 $f,g\in L^2(\mathbb{R})$,则成立 Plancherel 型公式

$$\int_{\mathbb{R}}f(t)\,\overline{g(t)}\mathrm{d}t=\frac{1}{2\pi}\int_{\mathbb{R}}f^{?}(\xi)\,\overline{g^{?}(\xi)}\mathrm{d}\xi,$$

若用 $L^2(\mathbb{R})$的内积表示,则得

$$(f,\overline{g})=\frac{1}{2\pi}(f^{?},\overline{g^{?}}).$$

注 在 Fourier 变换与逆变换的定义由对称形式给出时,Plancherel 公式为$(f,\overline{g})=(f^{?},\overline{g^{?}})$.

5. Fourier 变换的反演公式

定理 5.2.16 (1) 若 $f\in L^p(\mathbb{R})(1<p\leqslant 2)$,则成立 Fourier 变换的反演公式

$$\lim_{\rho\to+\infty}\left\|f(\circ)-\frac{1}{2\pi}\int_{(-\rho,\rho)}f^{?}(\xi)\mathrm{e}^{\mathrm{i}(\circ,\xi)}\mathrm{d}\xi\right\|_{L^p(\mathbb{R})}=0;$$

(2) 若 $f\in L^p(\mathbb{R})(1<p\leqslant 2)$,且 $f^{?}\in L^1(\mathbb{R})$,则成立 Fourier 变换的反演公式

$$\frac{1}{2\pi}\int_{\mathbb{R}}f^{?}(\xi)\mathrm{e}^{\mathrm{i}x\xi}\mathrm{d}\xi=f(x),\quad \text{a.e.}\,x\in\mathbb{R};$$

(3) 若 $f\in L^p(\mathbb{R})(1<p\leqslant 2)$,且 $g,g^{?}\in L^1(\mathbb{R})$,则成立卷积的 Fourier 变换的反演公式

$$(f*g)(x)=\frac{1}{2\pi}\int_{\mathbb{R}}f^{?}(\xi)g^{?}(\xi)\mathrm{e}^{\mathrm{i}x\xi}\mathrm{d}\xi,\quad x\in\mathbb{R}.$$

把 $L^1(\mathbb{R})$情形中的三个求和因子一般化,有望由 $f\in L^p(\mathbb{R})(1<p\leqslant 2)$的 Fourier 变换 $f^{?}\equiv F^p f\in L^q(\mathbb{R})$求得 f 自身.

定义 5.2.5(求和因子) 函数 $\theta(x)$称为求和因子,若满足

(1) $\theta(x)$是偶函数;

(2) $\theta(x)\in L^1(\mathbb{R})$, $\theta^{?}(\xi)\in L^1(\mathbb{R})$;

(3) $\int_{\mathbb{R}}\theta^{?}(\xi)\mathrm{d}\xi=2\pi$.

进而,若求和因子 $\theta(x)$是$\mathbb{R}$上的连续函数,则称 $\theta(x)$为连续的求和因子.

重要的求和因子就是我们已经遇到的三种:

(1) Cesaro 因子 $\theta(x)=\begin{cases}1-|x|, & |x|\leqslant 1,\\ 0, & |x|>1;\end{cases}$

(2) Abel 因子 $\theta(x)=\mathrm{e}^{-|x|}$, $x\in\mathbb{R}$;

(3) Gauss 因子 $\theta(x)=\mathrm{e}^{-x^2}$ $x\in\mathbb{R}$.

称

$$\frac{1}{2\pi}\int_{\mathbb{R}}f^{?}(\xi)\mathrm{e}^{\mathrm{i}x\xi}\mathrm{d}\xi,\quad x\in\mathbb{R}$$

为 **Fourier 反演积分**,而称

$$U(f,x,\rho)=\frac{1}{2\pi}\int_{\mathbb{R}}\theta\left(\frac{\xi}{\rho}\right)f^{?}(\xi)\mathrm{e}^{\mathrm{i}x\xi}\mathrm{d}\xi,\quad x\in\mathbb{R}$$

为 **Fourier 反演积分的 θ 和**.

定理 5.2.17 若 $f\in L^p(\mathbb{R})(1<p\leqslant 2)$,则对于求和因子 $\theta(\xi)$,Fourier 反演积分的 θ 和

$$U(f,x,\rho)=\frac{1}{2\pi}\int_{\mathbb{R}}\theta\left(\frac{\xi}{\rho}\right)f^{?}(\xi)\mathrm{e}^{\mathrm{i}x\xi}\mathrm{d}\xi$$

满足

$$\lim_{\rho\to+\infty}\|U(f,\circ,\rho)-f(\circ)\|_{L^p(\mathbb{R})}=0,$$

并且

$$\|U(f,\circ,\rho)\|_{L^p(\mathbb{R})}\leqslant\|\theta^{?}\|_{L^1(\mathbb{R})}\|f\|_{L^p(\mathbb{R})},\quad \rho>0;$$

此外,当求和因子 $\theta^{?}\geqslant 0$,并且 $\theta^{?}$在$[0,+\infty)$中单调递减时,则 Fourier 反演积分几乎处处 θ 收敛于 f,即

$$\lim_{\rho\to+\infty}U(f,x,\rho)=\lim_{\rho\to+\infty}\frac{1}{2\pi}\int_{\mathbb{R}}\theta\left(\frac{\xi}{\rho}\right)f^{?}(\xi)\mathrm{e}^{\mathrm{i}x\xi}\mathrm{d}\xi=f(x),\quad \text{a. e. } x\in\mathbb{R}.$$

事实上,上述定理对于 $f\in L^p(\mathbb{R})(1\leqslant p\leqslant 2)$都成立.

利用定理 5.2.17,可以证明定理 5.2.8 的反演公式对于 $L^p(\mathbb{R})(1\leqslant p\leqslant 2)$上的 Fourier 变换成立.

6. Fourier 变换的惟一性定理

定理 5.2.18 设 $f,g\in L^p(\mathbb{R})(1\leqslant p\leqslant 2)$,若 $f^?(\xi)=g^?(\xi)$,a. e.,则

$$f(x)=g(x),\quad \text{a. e. } x\in\mathbb{R}.$$

5.3 Schwartz 分布的 Fourier 变换

5.3.1 Schwartz 函数的 Fourier 变换

1. $\mathbb{S}(\mathbb{R}^n)$上的 Fourier 变换

在 5.1 节、5.2 节中,我们定义了 $f\in L^p(\mathbb{R}^n)(1\leqslant p\leqslant 2)$的 Fourier 变换,并研究了它们的性质. 对于 Schwartz 函数的 Fourier 变换、分布的 Fourier 变换,将在本节中讨论.

首先讨论 Schwartz 函数 $\varphi\in\mathbb{S}(\mathbb{R}^n)$的 Fourier 变换. 由于$\mathbb{S}(\mathbb{R}^n)\subset L^1(\mathbb{R}^n)$,故可以在 $L^1(\mathbb{R}^n)$中函数的 Fourier 变换的基础上,给出$\mathbb{S}(\mathbb{R}^n)$中函数的 Fourier 变换,并且这里需要在 $\mathbb{R}^n$ 上进行讨论,因为$\mathbb{S}(\mathbb{R}^n)$中的函数是定义在$\mathbb{R}^n$上的.

本节所要讨论的许多性质不一一给出证明,只是写成定理形式,或证明其中的一部分. 有兴趣或有需要的读者可以自己证明,或者从参考书中找到答案.

定义 5.3.1 函数 $\varphi\in\mathbb{S}(\mathbb{R}^n)$的 Fourier 变换定义为

$$\varphi^?(\xi)=\int_{\mathbb{R}^n}\varphi(x)\mathrm{e}^{-\mathrm{i}x\cdot\xi}\mathrm{d}x,\quad \xi\in\mathbb{R}^n,$$

其中 $x\cdot\xi=\sum\limits_{j=1}^{n}x_j\xi_j$ 是 $x=(x_1,x_2,\cdots,x_n)\in\mathbb{R}^n$ 与 $\xi=(\xi_1,\xi_2,\cdots,\xi_n)\in\mathbb{R}^n$ 的内积;(比较定义 5.2.1)

函数 $\psi\in\mathbb{S}(\mathbb{R}^n)$的 Fourier 逆变换定义为

$$\psi^?(x)=\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\psi(\xi)\mathrm{e}^{\mathrm{i}x\cdot\xi}\mathrm{d}\xi,\quad x\in\mathbb{R}^n,$$

其中 $x\cdot\xi=\sum\limits_{j=1}^{n}x_j\xi_j$ 是 $x=(x_1,x_2,\cdots,x_n)\in\mathbb{R}^n$ 与 $\xi=(\xi_1,\xi_2,\cdots,\xi_n)\in\mathbb{R}^n$ 的内积.(比较定义 5.2.2)

2. $\mathbb{S}(\mathbb{R}^n)$上的 Fourier 变换的性质

在$\mathbb{R}^n$ 情形使用以下记号.

对于 $x=(x_1,x_2,\cdots,x_n)\in\mathbb{R}^n$，$\beta=(\beta_1,\beta_2,\cdots,\beta_n)\in\mathbb{N}^n$，记 $|x|=\sqrt{\sum_{j=1}^{n}x_j^2}$ 与 $x^\beta=x_1^{\beta_1}x_2^{\beta_2}\cdots x_n^{\beta_n}$；对于 $\alpha=(\alpha_1,\alpha_2,\cdots,\alpha_n)\in\mathbb{N}^n$，记 $\alpha!=\alpha_1!\ \alpha_2!\ \cdots\alpha_n!$；记$\partial_j\equiv\partial_{x_j}=\dfrac{\partial}{\partial x_j}(j=1,2,\cdots,n)$与 $D_j=\dfrac{1}{\mathrm{i}}\partial_j(j=1,2,\cdots,n)$；记

$$D_x^\alpha\equiv D_1^{\alpha_1}\cdots D_n^{\alpha_n}=\left(\frac{1}{\mathrm{i}}\right)^{\alpha_1+\cdots+\alpha_n}\partial_{x_1}^{\alpha_1}\cdots\partial_{x_n}^{\alpha_n}$$

用这些记号会带来许多方便.

1）基本运算性质

定理 5.3.1 函数 $\varphi\in\mathbb{S}(\mathbb{R}^n)$的 Fourier 变换的基本运算性质如下：

(1) $(\tau_h\varphi(\circ))^?(\xi)=\mathrm{e}^{-ih\cdot\xi}\varphi^?(\xi)$，$\xi\in\mathbb{R}^n$，$h\in\mathbb{R}^n$； （平移的 FT）

(2) $\tau_h\varphi^?(\xi)=[\mathrm{e}^{ih\cdot\circ}\varphi(\circ)]^?(\xi)$，$\xi\in\mathbb{R}^n$，$h\in\mathbb{R}^n$； （FT 的平移）

(3) $[\varphi(\rho\circ)]^?(\xi)=|\rho|^{-n}\varphi^?(\rho^{-1}\xi)$，$\xi\in\mathbb{R}^n$，$\rho\neq0$； （伸缩的 FT）

(4) $\varphi^?(\rho\xi)=|\rho|^{-n}[\varphi(\rho^{-1}\circ)]^?(\xi)$，$\xi\in\mathbb{R}^n$，$\rho\neq0$； （FT 的伸缩）

(5) $(\tilde{\varphi}(\circ))^?(\xi)=(2\pi)^n\varphi^?(\xi)$，$\xi\in\mathbb{R}^n$； （反射的 FT）

(6) $(\varphi^?(\circ))^\sim(\xi)=(2\pi)^n\varphi^?(\xi)$，$\xi\in\mathbb{R}^n$； （FT 的反射）

(7) $(\bar{\varphi}(\circ))^?(\xi)=(2\pi)^n\ \overline{\varphi^?}(\xi)$，$\xi\in\mathbb{R}^n$； （共轭的 FT）

(8) $\overline{\varphi^?}(\xi)=(2\pi)^n\ (\bar{\varphi})^?(\xi)$，$\xi\in\mathbb{R}^n$； （FT 的共轭）

(9) $(D_x^p\varphi(\circ))^?(\xi)=\xi^p\varphi^?(\xi)$，$\xi\in\mathbb{R}^n$，$p\in\mathbb{N}^n$； （导数的 FT）

(10) $D_\xi^p\varphi^?(\xi)=((-\circ)^p\varphi(\circ))^?(\xi)$，$\xi\in\mathbb{R}^n$，$p\in\mathbb{N}^n$； （FT 的导数）

(11) $(\varphi*\psi(\circ))^?(\xi)=\varphi^?(\xi)\psi^?(\xi)$，$\xi\in\mathbb{R}^n$； （卷积的 FT）

(12) $(\varphi^?*\psi^?)(\xi)=(2\pi)^n(\varphi(\circ)\psi(\circ))^?(\xi)$，$\xi\in\mathbb{R}^n$. （FT 的卷积）

证 性质(1)～(4)可直接由定义得到(与定理 5.2.1 比较).

对于(5)，由$\mathbb{S}(\mathbb{R}^n)\subset L^1(\mathbb{R}^n)$，且 $\varphi\in\mathbb{S}(\mathbb{R}^n)\Rightarrow\tilde{\varphi}\in\mathbb{S}(\mathbb{R}^n)$，故 Fourier 变换$(\tilde{\varphi})^?$存在，

$$\begin{aligned}(\tilde{\varphi}(\circ))^?(\xi)&=\int_{\mathbb{R}^n}\varphi(-x)\mathrm{e}^{-\mathrm{i}x\cdot\xi}\mathrm{d}x=\int_{\mathbb{R}^n}\varphi(t)\mathrm{e}^{\mathrm{i}t\cdot\xi}\mathrm{d}t\\&=(2\pi)^n\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\varphi(t)\mathrm{e}^{\mathrm{i}t\cdot\xi}\mathrm{d}t=(2\pi)^n\varphi^?(\xi)\end{aligned}$$

得到.

注意到，在定理 5.2.1 中，反射的 FT 公式为$\left[\overline{f(-\circ)}\right]^?(\xi)=\overline{f^?}(\xi)(\xi\in\mathbb{R})$，是由

$$\begin{aligned}[\overline{f(-\circ)}]^?(\xi)&=(\overline{\tilde{f}(\circ)})^?(\xi)=\int_{\mathbb{R}^n}\overline{f(-x)}\mathrm{e}^{-\mathrm{i}x\cdot\xi}\mathrm{d}x=\int_{\mathbb{R}^n}\overline{f(t)}\mathrm{e}^{\mathrm{i}t\cdot\xi}\mathrm{d}t\\&=\int_{\mathbb{R}^n}\overline{f(t)\mathrm{e}^{-\mathrm{i}t\cdot\xi}}\mathrm{d}t=\overline{f^?}(\xi)\end{aligned}$$

得到. 只是反射的 FT 的不同表示而已.

性质(6)～(12)只对$\mathbb{R}$的情形证明,一般$\mathbb{R}^n$情形类似可证.

对于(6),由

$$(\varphi^{?}(\circ))^{\sim}(\xi)=\left(\int_{\mathbb{R}}\varphi(x)\mathrm{e}^{-\mathrm{i}x\xi}\mathrm{d}x\right)^{\sim}=\int_{\mathbb{R}}\varphi(x)\mathrm{e}^{\mathrm{i}x\xi}\mathrm{d}x$$
$$=(2\pi)^{1}\frac{1}{(2\pi)^{1}}\int_{\mathbb{R}}\varphi(x)\mathrm{e}^{\mathrm{i}x\xi}\mathrm{d}x=(2\pi)\varphi^{?}(\xi)$$

得到.

对于(7),由

$$(\overline{\varphi}(\circ))^{?}(\xi)=\int_{\mathbb{R}}\overline{\varphi}(x)\mathrm{e}^{-\mathrm{i}x\xi}\mathrm{d}x=\overline{\int_{\mathbb{R}}\varphi(x)\mathrm{e}^{\mathrm{i}x\xi}\mathrm{d}x}=$$
$$=(2\pi)^{1}\frac{1}{(2\pi)^{1}}\overline{\int_{\mathbb{R}}\varphi(x)\ \mathrm{e}^{\mathrm{i}x\xi}\mathrm{d}x}=(2\pi)\ \overline{\varphi^{?}}(\xi)$$

得到.

对于(8),由

$$\overline{\varphi^{?}}(\xi)=\overline{\int_{\mathbb{R}}\varphi(x)\mathrm{e}^{-\mathrm{i}x\xi}\mathrm{d}x}=\int_{\mathbb{R}}\overline{\varphi}(x)\mathrm{e}^{\mathrm{i}x\xi}\mathrm{d}x=(2\pi)^{1}\frac{1}{(2\pi)^{1}}\int_{\mathbb{R}}\overline{\varphi}(x)\mathrm{e}^{\mathrm{i}x\xi}\mathrm{d}x=(2\pi)(\overline{\varphi})^{?}(\xi)$$

得到.

对于(9),由 $\varphi^{?}(\xi)=\int_{\mathbb{R}}\varphi(x)\mathrm{e}^{-\mathrm{i}x\xi}\mathrm{d}x$ 给出

$$\xi^{p}\varphi^{?}(\xi)=\int_{\mathbb{R}}\xi^{p}\varphi(x)\mathrm{e}^{-\mathrm{i}x\xi}\mathrm{d}x=\int_{\mathbb{R}}\varphi(x)\{(-D_x)^{p}(\mathrm{e}^{-\mathrm{i}x\xi})\}\mathrm{d}x \qquad \text{(分部积分)}$$
$$=\{\varphi(x)(\mathrm{e}^{-\mathrm{i}x\xi})\}\Big|_{-\infty}^{+\infty}-\int_{\mathbb{R}}\{(-D_x)^{p}\varphi(x)\}\mathrm{e}^{-\mathrm{i}x\xi}\mathrm{d}x=\int_{\mathbb{R}}\{(D_x^{p})\varphi(x)\}\mathrm{e}^{-\mathrm{i}x\xi}\mathrm{d}x,$$

此即(9)的结论.

对于(10),由$\varphi\in\mathbb{S}(\mathbb{R})$所满足的条件,对$\varphi^{?}(\xi)=\int_{\mathbb{R}}\varphi(x)\mathrm{e}^{-\mathrm{i}x\xi}\mathrm{d}x$可以积分号下求导数,得

$$D_{\xi}^{p}\varphi^{?}(\xi)=\int_{\mathbb{R}}(-x)^{p}\varphi(x)\mathrm{e}^{-\mathrm{i}x\xi}\mathrm{d}x;$$

此即(10) 的结论.

由(9)、(10)还可立即推出$\varphi\in\mathbb{S}(\mathbb{R})$蕴含$\varphi^{?}\in\mathbb{S}(\mathbb{R})$. 事实上,由$\varphi\in\mathbb{S}(\mathbb{R})$,知积分

$$|\xi^{\beta}\varphi^{?}(\xi)|=\left|\int_{\mathbb{R}}\{(\mathrm{D}_x^{\beta})\varphi(x)\}\mathrm{e}^{-\mathrm{i}x\xi}\mathrm{d}x\right|\leqslant\int_{\mathbb{R}}|(\mathrm{D}_x^{\beta})\varphi(x)|\mathrm{d}x$$

与

$$|\xi^{\beta}\mathrm{D}_{\xi}^{p}\varphi^{?}(\xi)|=\left|\int_{\mathbb{R}}\mathrm{e}^{-\mathrm{i}x\xi}\mathrm{D}_x^{p}[(-x)^{\beta}\varphi(x)]\mathrm{d}x\right|$$

$$\leqslant \left|\int_{\mathbb{R}}(1+|x|^2)^{-\frac{n+1}{2}}(1+|x|^2)^{\frac{n+1}{2}}D_x^p[(-x)^\beta\varphi(x)]dx\right|$$

$$\leqslant c\sup_{\mathbb{R}}\left|(1+|x|^2)^{\frac{n+1}{2}}D_x^p[(-x)^\beta\varphi(x)]\right|$$

都是有界的,因此 $\varphi^? \in \mathbb{S}(\mathbb{R})$. 于是,$\varphi\in\mathbb{S}(\mathbb{R}^n)\Rightarrow\varphi^?\in\mathbb{S}(\mathbb{R}^n)\Rightarrow(\varphi^?)^?=\varphi=(\varphi^?)^?$.

对于(11),由 Fubini 定理,得到

$$(\varphi*\psi(\circ))^?(\xi)=\int_{\mathbb{R}}(\varphi*\psi)(x)e^{-ix\xi}dx=\int_{\mathbb{R}}\int_{\mathbb{R}}e^{-i(x-y)\xi}\varphi(x-y)\psi(y)e^{-iy\xi}dydx$$

$$=\int_{\mathbb{R}}e^{-ix\xi}\varphi(x)dx\cdot\int_{\mathbb{R}}e^{-iy\xi}\psi(y)dy=\varphi^?(\xi)\psi^?(\xi).$$

同时不难得出 $(\varphi*\psi(\circ))^?(\xi)=2\pi\varphi^?(\xi)\psi^?(\xi)$. 由此还说明

$$\varphi,\psi\in\mathbb{S}(\mathbb{R})\Rightarrow\varphi^?\psi^?\in\mathbb{S}(\mathbb{R})\Rightarrow(\varphi*\psi)^?\in\mathbb{S}(\mathbb{R})\Rightarrow\varphi*\psi\in\mathbb{S}(\mathbb{R}).$$

对于(12),由 Fourier 变换的惟一性定理与(11),对(12)式左边作 Fourier 逆变换,得

$$[(\varphi^?*\psi^?)(\circ)]^?(x)=2\pi(\varphi^?)^?(x)\cdot(\psi^?)^?(x)=2\pi\varphi(x)\psi(x);$$

对(12)式右边作 Fourier 逆变换,得

$$\{[(2\pi)(\varphi(\circ)\cdot\psi(\circ)]^?(\circ)\}^?(x)=(2\pi)[\varphi(x)\cdot\psi(x)]=2\pi\varphi(x)\psi(x).$$

因此(12)成立. 定理得证.

2) Fourier 变换的分析性质

定理 5.3.2 函数 $\varphi\in\mathbb{S}(\mathbb{R}^n)$ 的 Fourier 变换 $F\varphi\equiv\varphi^?$ 具有下述性质:

(1) $\mathbb{S}(\mathbb{R}^n)$ 中的 Riemann-Lebesgue 引理: $\lim\limits_{|\xi|\to+\infty}\varphi^?(\xi)=0,\forall\varphi\in\mathbb{S}(\mathbb{R}^n)$;

(2) $F:\varphi\to\varphi^?$ 是 $\mathbb{S}(\mathbb{R}^n)$ 到 $\mathbb{S}(\mathbb{R}^n)$ 上的连续线性算子;

(3) $F:\mathbb{S}(\mathbb{R}^n)\to\mathbb{S}(\mathbb{R}^n)$ 有连续逆算子 $F^{-1}:\mathbb{S}(\mathbb{R}^n)\to\mathbb{S}(\mathbb{R}^n)$,并且对 Fourier 变换 $F:\varphi\to\varphi^?$ 的逆变换 $F^{-1}:\varphi^?\to\varphi$,在 $\mathbb{S}(\mathbb{R}^n)$ 上成立 Fourier 变换的反演公式

$$\varphi(x)=\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\varphi^?(\xi)e^{i\xi\cdot x}d\xi;$$

此即 $(\varphi^?(\circ))^?(x)=\varphi(x)=(\varphi^?(\circ))^?(x)$;从而 $F:\mathbb{S}(\mathbb{R}^n)\to\mathbb{S}(\mathbb{R}^n)$ 是 $\mathbb{S}(\mathbb{R}^n)$ 到 $\mathbb{S}(\mathbb{R}^n)$ 的拓扑同构;

(4) $\mathbb{S}(\mathbb{R}^n)$ 上的 Parseval 公式:

$$\int_{\mathbb{R}^n}\varphi(x)\psi^?(x)dx=\int_{\mathbb{R}^n}\varphi^?(x)\psi(x)dx;\quad\left(\langle\varphi,\psi^?\rangle=\langle\varphi^?,\psi\rangle\right)$$

Plancherel 型公式:

$$\int_{\mathbb{R}^n}\varphi(x)\bar{\psi}(x)dx=\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\varphi^?(\xi)\overline{\psi^?}(\xi)d\xi.\quad\left(\langle\varphi,\bar{\psi}\rangle=\frac{1}{(2\pi)^n}\langle\varphi^?,\overline{\psi^?}\rangle\right)$$

证 只对 $\mathbb{R}$ 情形证明. 对于(1),可由 $L^1(\mathbb{R})$ 的 Riemann-Legesgue 引理得到. (2)是显

然的. 为证(3),利用 Gauss 函数 $g(x)=e^{-\frac{|x|^2}{2}}$ 的 Fourier 变换

$$g^?(\xi)=\int_{\mathbb{R}}e^{-\frac{|x|^2}{2}}e^{-ix\xi}dx=e^{-\frac{|\xi|^2}{2}}\int_{\mathbb{R}}e^{-\frac{1}{2}(x+i\xi)^2}dx=\sqrt{2\pi}e^{-\frac{|\xi|^2}{2}}$$

$\left(\text{由已知的欧拉积分}\int_{\mathbb{R}}e^{-\frac{t^2}{2}}dt=\sqrt{2\pi}\text{ 得到}\right)$,由

$$\begin{aligned}\int_{\mathbb{R}}\varphi^?(\xi)g(\xi)e^{ix\xi}d\xi&=\int_{\mathbb{R}}g(\xi)e^{ix\xi}d\xi\int_{\mathbb{R}}\varphi(y)e^{-iy\xi}dy\\&=\int_{\mathbb{R}}\int_{\mathbb{R}}e^{i(x-y)\xi}g(\xi)\varphi(y)d\xi dy=\int_{\mathbb{R}}g^?(y-x)\varphi(y)dy\\&=\int_{\mathbb{R}}g^?(y)\varphi(x+y)dy,\end{aligned}$$

用 $g(\varepsilon\xi)(\varepsilon>0)$ 代替上式中的 $g(\xi)$,则 $g^?(y)$ 换为 $\varepsilon^{-1}g^?\left(\frac{y}{\varepsilon}\right)$;再令 $\frac{y}{\varepsilon}=y_1$,则上式化为

$$\int_{\mathbb{R}}\varphi^?(\xi)g(\varepsilon\xi)e^{ix\xi}d\xi=\int_{\mathbb{R}}\varepsilon^{-1}g^?\left(\frac{y}{\varepsilon}\right)\varphi(x+y)dy=\int_{\mathbb{R}}g^?(y_1)\varphi(x+\varepsilon y_1)dy_1;$$

令 $\varepsilon\to0$,得到

$$g(0)\int_{\mathbb{R}}\varphi^?(\xi)e^{ix\xi}d\xi=\varphi(x)\int_{\mathbb{R}}g^?(y_1)dy_1.$$

但是,$g(0)=1,\int_{\mathbb{R}}g^?(y_1)dy_1=\sqrt{2\pi}\int_{\mathbb{R}}e^{-\frac{|y_1|^2}{2}}dy_1=2\pi$,代入上式,得到

$$\varphi(x)=\frac{1}{2\pi}\int_{\mathbb{R}}\varphi^?(\xi)e^{ix\xi}d\xi.$$

F^{-1} 的连续性由此反演公式得到.

最后,(4)的 Parseval 公式由 Fubini 定理得到. Plancherel 型公式的证明较为复杂,我们省略. 读者可参看[4].

5.3.2 Schwartz 分布的 Fourier 变换

1. $\mathbb{S}^*(\mathbb{R}^n)$ 上的 Fourier 变换

定义 5.3.2 分布 $T\in\mathbb{S}^*(\mathbb{R}^n)$ 的 Fourier 变换 $T^?$ 定义为一个分布,它满足

$$\langle T^?,\varphi\rangle=\langle T,\varphi^?\rangle,\quad\forall\varphi\in\mathbb{S}(\mathbb{R}^n);$$

或等价地,有

$$\langle T^?,\varphi^?\rangle=\langle T,\varphi\rangle,\quad\forall\varphi\in\mathbb{S}(\mathbb{R}^n).$$

分布 $T\in\mathbb{S}^*(\mathbb{R}^n)$ 的 Fourier 逆变换 $T^?$ 定义为一个分布,满足

$$\langle T^?,\varphi\rangle=\langle T,\varphi^?\rangle,\quad\forall\varphi\in\mathbb{S}(\mathbb{R}^n);$$

或等价地,有

$$\langle T^?,\varphi^?\rangle=\langle T,\varphi\rangle,\quad\forall\varphi\in\mathbb{S}(\mathbb{R}^n).$$

2. $\mathbb{S}^*(\mathbb{R}^n)$上的 Fourier 变换的性质

1）基本运算性质

Schwartz 分布的 Fourier 变换的运算性质：

定理 5.3.3 分布 $T\in\mathbb{S}^*(\mathbb{R}^n)$的 Fourier 变换的基本运算性质如下：

(1) $(\tau_h T)^{?}=\mathrm{e}^{-\mathrm{i}h\cdot\xi}T^{?}$，$h\in\mathbb{R}^n$，$\xi\in\mathbb{R}^n$； （平移的 FT）

(2) $\tau_h T^{?}=(\mathrm{e}^{\mathrm{i}h\cdot\circ}T)^{?}$，$h\in\mathbb{R}^n$； （FT 的平移）

(3) $(\widetilde{T})^{?}=(2\pi)^n T^{?}$； （反射的 FT）

(4) $(T^{?})^{\sim}=(2\pi)^n T^{?}$； （FT 的反射）

(5) $(D^p T)^{?}=\xi^p T^{?}$，$\xi\in\mathbb{R}^n$，$p\in\mathbb{N}^n$； （导数的 FT）

(6) $(-D)^p T^{?}=(\circ^p T)^{?}$，$p\in\mathbb{N}^n$. （FT 的导数）

证 为证(1)，任取 $\varphi\in\mathbb{S}(\mathbb{R}^n)$，故 $\forall h\in\mathbb{R}^n$，(1)左边$(\tau_h T)^{?}$对 $\varphi\in\mathbb{S}(\mathbb{R}^n)$的作用为

$$\langle(\tau_h T)^{?},\varphi\rangle=\langle\tau_h T,\varphi^{?}\rangle=\langle T,\tau_{-h}\varphi^{?}\rangle;$$

式(1)右边 $\mathrm{e}^{-\mathrm{i}h\cdot\xi}T^{?}$对 $\varphi\in\mathbb{S}(\mathbb{R}^n)$的作用为

$$\langle\mathrm{e}^{-\mathrm{i}h\circ\xi}T^{?},\varphi\rangle=\langle T^{?},\mathrm{e}^{-\mathrm{i}h\circ\xi}\varphi\rangle=\langle T,(\mathrm{e}^{-\mathrm{i}h\cdot\circ}\varphi)^{?}\rangle=\langle T,\tau_{-h}\varphi^{?}\rangle,$$

故左右相等，于是得到$(\tau_h T)^{?}=\mathrm{e}^{-\mathrm{i}h\cdot\xi}T^{?}$，$(h,\xi\in\mathbb{R}^n)$. 式(1)得证.

为证(2)，任取 $\varphi\in\mathbb{S}(\mathbb{R}^n)$，于是，$\forall h\in\mathbb{R}^n$，(2)左边 $\tau_h T^{?}$ 对 $\varphi\in\mathbb{S}(\mathbb{R}^n)$的作用为

$$\langle\tau_h T^{?},\varphi\rangle=\langle T^{?},\tau_{-h}\varphi\rangle=\langle T,(\tau_{-h}\varphi)^{?}\rangle=\langle T,\mathrm{e}^{\mathrm{i}h\cdot\xi}\varphi^{?}\rangle=\langle\mathrm{e}^{\mathrm{i}h\cdot\xi}T,\varphi^{?}\rangle=\langle(\mathrm{e}^{\mathrm{i}h\cdot\circ}T)^{?},\varphi\rangle,$$

故 $\tau_h T^{?}=(\mathrm{e}^{\mathrm{i}h\cdot\circ}T)^{?}$，$(h\in\mathbb{R}^n)$，式(2)得证.

为证(3)，任取 $\varphi\in\mathbb{S}(\mathbb{R}^n)$，于是，(3)左边$(\widetilde{T})^{?}$对 $\varphi\in\mathbb{S}(\mathbb{R}^n)$的作用为

$$\langle(\widetilde{T})^{?},\varphi\rangle=\langle\widetilde{T},\varphi^{?}\rangle=\langle T,\widetilde{\varphi^{?}}\rangle=\langle T,(2\pi)^n\varphi^{?}\rangle=\langle(2\pi)^n T,\varphi^{?}\rangle;$$

式(3)右边$(2\pi)^n T^{?}$ 对 $\varphi\in\mathbb{S}(\mathbb{R}^n)$的作用为$\langle(2\pi)^n T^{?},\varphi\rangle=\langle(2\pi)^n T,\varphi^{?}\rangle$，故得到$(\widetilde{T})^{?}=(2\pi)^n T^{?}$，式(3)得证.

为证(4)，由

$$\langle(T^{?})^{\sim},\varphi\rangle=\langle T^{?},\widetilde{\varphi}\rangle=\langle T,(\widetilde{\varphi})^{?}\rangle=\langle T,(2\pi)^n\varphi^{?}\rangle=\langle(2\pi)^n T^{?},\varphi\rangle,$$

即得(4).

为证(5)，任取 $\varphi\in\mathbb{S}(\mathbb{R}^n)$，于是，对于任取的 $p=\alpha\in\mathbb{N}^n$，(5)左边$(D^\alpha T)^{?}$对 $\varphi\in\mathbb{S}(\mathbb{R}^n)$的作用为

$$\begin{aligned}\langle(D^\alpha T)^{?},\varphi\rangle&=\langle D^\alpha T,\varphi^{?}\rangle=(-1)^{|\alpha|}\langle T,D^\alpha\varphi^{?}\rangle=(-1)^{|\alpha|}\langle T,(-1)^{|\alpha|}(-D)^\alpha\varphi^{?}\rangle\\&=(-1)^{2|\alpha|}\langle T,(-D)^\alpha\varphi^{?}\rangle=\langle T,(-D)^\alpha\varphi^{?}\rangle=\langle T,(\circ^\alpha\varphi)^{?}\rangle;\end{aligned}$$

式(5)右边 $\xi^\alpha T^{?}$ 对 $\varphi\in\mathbb{S}(\mathbb{R}^n)$的作用为$\langle\xi^\alpha T^{?},\varphi\rangle=\langle T^{?},\xi^\alpha\varphi\rangle=\langle T,(\circ^\alpha\varphi)^{?}\rangle$，于是

$$(D^\alpha T)^{?}=\xi^\alpha T^{?},\quad\xi\in\mathbb{R}^n.$$

即得(5).

为证(6),取 $\varphi\in\mathbb{S}(\mathbb{R}^n)$,对于任取的 $p=\alpha\in\mathbb{N}^n$,则由定理 5.3.1(8),得

$$\begin{aligned}\langle(\circ^{\alpha}T)^{?},\varphi\rangle &= \langle x^{\alpha}T,\varphi^{?}\rangle = \langle T,\xi^{\alpha}\varphi^{?}\rangle = \langle T,(D^{\alpha}\varphi)^{?}\rangle \\ &= \langle T^{?},D^{\alpha}\varphi\rangle = \langle(-D)^{\alpha}T^{?},\varphi\rangle,\end{aligned}$$

故 $(-D)^{\alpha}T^{?}=(\circ^{\alpha}T)^{?}$,(6)得证.

2) 分析性质

Schwartz 分布 Fourier 变换的分析性质可归纳为如下定理.

定理 5.3.4 分布 $T\in\mathbb{S}^*(\mathbb{R}^n)$的 Fourier 变换 $T^{?}$ 具有下述性质:

(1) $F:T\to T^{?}$ 是$\mathbb{S}^*(\mathbb{R}^n)$到$\mathbb{S}^*(\mathbb{R}^n)$上的连续线性算子(在$\mathbb{S}^*(\mathbb{R}^n)$的强、弱*拓扑下);

(2) $F:\mathbb{S}^*(\mathbb{R}^n)\to\mathbb{S}^*(\mathbb{R}^n)$有连续逆算子 $F^{-1}:\mathbb{S}^*(\mathbb{R}^n)\to\mathbb{S}^*(\mathbb{R}^n)$,并且对于 $F:\varphi\to\varphi^{?}$ 的逆变换 $F^{-1}:\varphi^{?}\to\varphi$,在$\mathbb{S}^*(\mathbb{R}^n)$上,有

$$(T^{?})^{?} = T \quad 与 \quad (T^{?})^{?} = T.$$

由此得到,$F:\mathbb{S}^*(\mathbb{R}^n)\to\mathbb{S}^*(\mathbb{R}^n)$是$\mathbb{S}^*(\mathbb{R}^n)$到$\mathbb{S}^*(\mathbb{R}^n)$的拓扑同构.

证 对于(1),$F:\mathbb{S}^*(\mathbb{R}^n)\to\mathbb{S}^*(\mathbb{R}^n)$的线性由定义得到.

对于 $F:\mathbb{S}^*(\mathbb{R}^n)\to\mathbb{S}^*(\mathbb{R}^n)$的连续性,由于 $F:\mathbb{S}(\mathbb{R}^n)\to\mathbb{S}(\mathbb{R}^n)$是空间$\mathbb{S}(\mathbb{R}^n)$到空间$\mathbb{S}(\mathbb{R}^n)$的连续线性算子(定理 5.3.2),故其共轭空间$\mathbb{S}^*(\mathbb{R}^n)$到自身$\mathbb{S}^*(\mathbb{R}^n)$的算子 $F:\mathbb{S}^*(\mathbb{R}^n)\to\mathbb{S}^*(\mathbb{R}^n)$在$\mathbb{S}^*(\mathbb{R}^n)$的强拓扑和弱*拓扑意义下也是连续的.

对于(2),只需证明 $F^{-1}:\mathbb{S}^*(\mathbb{R}^n)\to\mathbb{S}^*(\mathbb{R}^n)$是连续线性映射即可. 事实上,$\forall\, T\in\mathbb{S}^*(\mathbb{R}^n)$,由定理 5.3.2 知 $F:\mathbb{S}(\mathbb{R}^n)\to\mathbb{S}(\mathbb{R}^n)$是拓扑同构,$\langle T,\varphi\rangle(\varphi\in\mathbb{S}(\mathbb{R}^n))$也应由 $\varphi^{?}$ 决定,而且是空间 $\{\varphi^{?}:\varphi\in\mathbb{S}(\mathbb{R}^n)\}$上的连续线性泛函 g. 但是,$\{\varphi^{?}:\varphi\in\mathbb{S}(\mathbb{R}^n)\}=\mathbb{S}(\mathbb{R}^n)$,故 g 就是$\mathbb{S}(\mathbb{R}^n)$上的连续线性泛函,且$\langle T,\varphi\rangle=\langle g,\varphi^{?}\rangle$. 于是,由定义立得 $T^{?}=g$,这就证明了 F^{-1}有意义,且成立$\langle T,\varphi\rangle=\langle F^{-1}T,\varphi^{?}\rangle$. 由此也得到$(T^{?})^{?}=T$ 与$(T^{?})^{?}=T$.

F^{-1}的线性显然;连续性也可由$\langle F^{-1}T,\varphi^{?}\rangle=\langle T,\varphi\rangle$立即得到. 从而得到 Schwartz 分布的 Fourier 变换

$$F:\mathbb{S}^*(\mathbb{R}^n)\to\mathbb{S}^*(\mathbb{R}^n)$$

是$\mathbb{S}^*(\mathbb{R}^n)$到$\mathbb{S}^*(\mathbb{R}^n)$的拓扑同构.

定理 5.3.5 对于分布 $T\in\mathbb{S}^*(\mathbb{R}^n)$与函数 $g\in\mathbb{S}(\mathbb{R}^n)$,有

(1) $\langle T,\overline{g}\rangle=\dfrac{1}{(2\pi)^n}\langle T^{?},\overline{g^{?}}\rangle$;

(2) $\langle T,g\rangle=\langle T^{?},g^{?}\rangle$.

证 对于(1),为证 $\langle T,\overline{g}\rangle=\dfrac{1}{(2\pi)^n}\langle T^{?},\overline{g^{?}}\rangle$,取 $g\in\mathbb{S}(\mathbb{R}^n)$,将其共轭 $\overline{g}$改写为

$$\overline{g} = [(\overline{g})^{?}]^{?} = \left[\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\overline{g}(\xi)\mathrm{e}^{\mathrm{i}x\cdot\xi}\mathrm{d}\xi\right]^{?} \equiv h^{?},$$

于是,

$$\langle T,\overline{g}\rangle = \langle T,h^?\rangle = \langle T^?,h\rangle = \left\langle T^?,\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\overline{g}(\xi)\mathrm{e}^{\mathrm{i}x\cdot\xi}\mathrm{d}\xi\right\rangle$$

$$= \left\langle T^?,\frac{1}{(2\pi)^n}\overline{\int_{\mathbb{R}^n}g(\xi)\mathrm{e}^{-\mathrm{i}x\cdot\xi}\mathrm{d}\xi}\right\rangle = \frac{1}{(2\pi)^n}\left\langle T^?,\overline{g^?}\right\rangle.$$

对于(2),为证 $\langle T,g\rangle=\langle T^?,g^?\rangle$,由 $\langle T^?,g^?\rangle=\langle T,(g^?)^?\rangle=\langle T,g\rangle$便得.

5.3.3 具紧致支集的 Schwartz 分布

1. $E^*(\mathbb{R}^n)$分布与$\mathbb{S}^*(\mathbb{R}^n)$分布的例子

例 5.3.1 $L^p(\mathbb{R}^n)\subset\mathbb{S}^*(\mathbb{R}^n)(1\leqslant p<+\infty)$

在例 5.1.3 中已经给出 $L^1(\mathbb{R}^n)\subset\mathbb{S}^*(\mathbb{R}^n)$,本例是$\forall f\in L^p(\mathbb{R}^n)(1<p<+\infty)$,令

$$\langle f,\varphi\rangle = \int_{\mathbb{R}^n}f(x)\varphi(x)\mathrm{d}x,\quad \forall\varphi\in\mathbb{S}(\mathbb{R}^n),$$

此式定义了$\mathbb{S}(\mathbb{R}^n)$上的一个线性泛函,由 Hölder 不等式

$$|\langle f,\varphi\rangle| = \left|\int_{\mathbb{R}^n}f(x)\varphi(x)\mathrm{d}x\right|\leqslant\|f\|_{L^p(\mathbb{R}^n)}\|\varphi\|_{L^q(\mathbb{R}^n)},\quad \forall\varphi\in\mathbb{S}(\mathbb{R}^n)\subset L^q(\mathbb{R}^n)$$

得到它的连续性.因此得到 $L^p(\mathbb{R}^n)\subset\mathbb{S}^*(\mathbb{R}^n)$,$1<p<+\infty$.

下面只要证明 $\varphi\in\mathbb{S}(\mathbb{R}^n)\Rightarrow\varphi\in L^q(\mathbb{R}^n)(1<q<+\infty)$.

$$\|\varphi\|_{L^q(\mathbb{R}^n)} = \left(\int_{\mathbb{R}^n}|\varphi(x)|^q\mathrm{d}x\right)^{\frac{1}{q}}\leqslant\sup_{\mathbb{R}^n}(1+|x|^2)^{\frac{n+1}{2q}}|\varphi(x)|\left[\int_{\mathbb{R}^n}(1+|x|^2)^{-\frac{n+1}{2}}\mathrm{d}x\right]^{\frac{1}{q}}$$

$$= c\sup_{\mathbb{R}^n}(1+|x|^2)^{\frac{n+1}{2q}}|\varphi(x)|\leqslant c_1<+\infty,$$

故得 $\varphi\in\mathbb{S}(\mathbb{R}^n)\Rightarrow\varphi\in L^q(\mathbb{R}^n)$,并且$\forall g\in L^q(\mathbb{R}^n)\subset\mathbb{S}^*(\mathbb{R}^n)$对于 $\varphi\in\mathbb{S}(\mathbb{R}^n)$的作用满足

$$|\langle g,\varphi\rangle|\leqslant c_1\sup_{\mathbb{R}^n}(1+|x|^2)^{\frac{n+1}{2q}}|\varphi(x)|.$$

这样,$\forall f\in L^p(\mathbb{R}^n)(1\leqslant p<+\infty)$是一个正则 Schwartz 分布.此分布显然未必有紧致支集.

例 5.3.2 任一个常系数多项式都在$\mathbb{S}^*(\mathbb{R}^n)$中.

只需考虑单项式 x^m(m 为确定的自然数),则

$$\langle x^m,\varphi\rangle = \int_{\mathbb{R}^n}x^m\varphi(x)\mathrm{d}x,\quad \forall\varphi\in\mathbb{S}(\mathbb{R}^n),$$

由于 $\lim\limits_{|x|\to+\infty}x^s\cdot x^m\varphi(x)=0$,$\forall s\in\mathbb{N}^n$,故积分存在,且是线性、连续的.但此分布显然不具有紧致支集.

例 5.3.3 狄拉克(Dirac)分布 $\delta\in E^*(\mathbb{R}^n)\subset\mathbb{S}^*(\mathbb{R}^n)$.

显然,狄拉克分布 δ 是具有紧致支集的 Schwartz 分布

$$\langle\delta,\varphi\rangle = \varphi(0),\quad \forall\varphi\in\mathbb{S}(\mathbb{R}^n),$$

且 $\mathrm{supp}\delta=\{0\}$.

狄拉克分布 δ 的平移 $\tau_{t_0}\delta$ 为

$$\langle \tau_{t_0}\delta,\varphi\rangle = \langle \delta,\tau_{-t_0}\varphi\rangle = \langle \delta,\varphi(t+t_0)\rangle = \varphi(t+t_0)\big|_{t=0} = \varphi(t_0),\quad \forall\,\varphi\in \mathbb{S}(\mathbb{R}^n),$$

因此,$\tau_{t_0}\delta\in \mathbb{S}^*(\mathbb{R}^n)$.由 $\mathrm{supp}\tau_{t_0}\delta=\{t_0\}$,知 $\tau_{t_0}\delta\in E^*(\mathbb{R}^n)$是具有紧致支集的奇异分布.

例 5.3.4 狄拉克分布 δ 的 Fourier 变换.

求狄拉克分布 δ 的 Fourier 变换,由

$$\langle \delta^{?},\varphi\rangle = \langle \delta,\varphi^{?}\rangle = \varphi^{?}(0) = \int_{\mathbb{R}^n}\varphi(x)\mathrm{e}^{-\mathrm{i}x\cdot 0}\mathrm{d}x = \langle 1,\varphi\rangle,\quad \forall\,\varphi\in \mathbb{S}(\mathbb{R}^n),$$

于是,$\delta^{?}=1$.$\delta^{?}\in L^1_{\mathrm{loc}}(\mathbb{R}^n)$是一个正则分布,但不具有紧致支集.

例 5.3.5 支撑在 t_0 的狄拉克分布 δ_{t_0} 的 Fourier 变换.

支撑在 t_0 的狄拉克分布 δ_{t_0} 定义为 $\tau_{t_0}\delta$(例 5.3.3)

$$\langle \delta_{t_0},\varphi\rangle \equiv \langle \tau_{t_0}\delta,\varphi\rangle = \langle \delta,\tau_{-t_0}\varphi\rangle = \langle \delta,\varphi(\xi+t_0)\rangle = \varphi(t_0),\quad \forall\,\varphi\in \mathbb{S}(\mathbb{R}^n).$$

于是,由定理 5.3.3(2),有

$$(\delta_{t_0})^{?} = (\tau_{t_0}\delta)^{?} = \mathrm{e}^{-\mathrm{i}t_0\xi}\delta^{?} = \mathrm{e}^{-\mathrm{i}t_0\xi},\quad \xi\in \mathbb{R}^n.$$

因此,$\delta_{t_0}\in E^*(\mathbb{R}^n)$的 Fourier 变换为 $\delta_{t_0}^{?}\in \mathbb{S}^*(\mathbb{R}^n)$.显然,$\delta_{t_0}^{?}$ 是不具有紧致支集的正则分布.

例 5.3.6 1 的 Fourier 变换.

由定义知

$$\begin{aligned}\langle 1^{?},\varphi\rangle &= \langle 1,\varphi^{?}\rangle = \int_{\mathbb{R}^n}\varphi^{?}(\xi)\mathrm{d}\xi = \int_{\mathbb{R}^n}\varphi^{?}(\xi)\mathrm{e}^{\mathrm{i}(0\cdot\xi)}\mathrm{d}\xi\\ &= (2\pi)^n\frac{1}{(2\pi)^n}\int_{\mathbb{R}^n}\varphi^{?}(\xi)\mathrm{e}^{-\mathrm{i}(0\cdot\xi)}\mathrm{d}\xi = (2\pi)^n(\varphi^{?})^{?}(0)\\ &= (2\pi)^n\varphi(0) = (2\pi)^n\langle\delta,\varphi\rangle,\quad \forall\,\varphi\in \mathbb{S}(\mathbb{R}^n),\end{aligned}$$

从而得到 $1^{?}=(2\pi)^n\delta\in E^*(\mathbb{R}^n)$,它是一个具有紧致支集的奇异分布.

在一维情形,$1^{?}=2\pi\delta$.

2. 具紧致支集的 $E^*(\mathbb{R}^n)$分布与$\mathbb{S}^*(\mathbb{R}^n)$分布的构造

分布支集的定义,已在定义 5.1.4 给出.对于与 Schwartz 空间$\mathbb{S}(\mathbb{R}^n)$相关的函数空间 $D(\mathbb{R}^n)\equiv C_C^\infty(\mathbb{R}^n)$,$C^\infty(\mathbb{R}^n)\equiv E(\mathbb{R}^n)$有包含关系 $C_C^\infty(\mathbb{R}^n)\subset\mathbb{S}(\mathbb{R}^n)\subset C^\infty(\mathbb{R}^n)$,对应的分布空间的包含关系为

$$D^*(\mathbb{R}^n)\supset \mathbb{S}^*(\mathbb{R}^n)\supset E^*(\mathbb{R}^n).$$

定理 5.3.6($E^*(\mathbb{R}^n)$中分布的构造定理) 对于分布 $T\in E^*(\mathbb{R}^n)$,有

(1) $E^*(\mathbb{R}^n)=\{S\in D^*(\mathbb{R}^n):\mathrm{supp}S$ 是紧致集$\}$;

(2) $S\in E^*(\mathbb{R}^n)\Leftrightarrow S=\sum\limits_{|p|\leqslant r}\partial^p f,\ f\in C_C^r(\mathbb{R}^n),\ p\in\mathbb{N}^n,\ |p|=p_1+p_2+\cdots+p_n$,且 $\mathrm{supp}f\subset(\mathrm{supp}S)_\varepsilon$,这里$(\mathrm{supp}S)_\varepsilon$ 是 $\mathrm{supp}S$ 的 ε 邻域,$r\in\mathbb{N}$;

(3) 若$\mathrm{supp}S=\{0\}$,则 $S=\sum\limits_{|p|\leqslant k}c_p\delta^{(p)}$;亦即支集为$\{0\}$的分布 S 必可表示为狄拉克分

布 $\delta\in E^*(\mathbb{R}^n)$及其导数 $\delta^{(p)}(p\in\mathbb{N}^n)$的有限线性组合,$k\in\mathbb{N}$,$c_p\in\mathbb{R}$.

定理 5.3.7(**$\mathbb{S}^*(\mathbb{R}^n)$中分布的构造定理**) 对于分布 $T\in\mathbb{S}^*(\mathbb{R}^n)$,有

$T\in\mathbb{S}^*(\mathbb{R}^n)\Leftrightarrow T=\partial^p[(1+|x|^2)^{\frac{k}{2}}f(x)]$, $f\in C(\mathbb{R}^n)\cap B(\mathbb{R}^n)$, $p\in\mathbb{N}^n,k\in\mathbb{N}$,其中 $B(\mathbb{R}^n)$为$\mathbb{R}^n$ 上的有界函数类.

以上两个定理的证明参看[5].

5.3.4 Schwartz 分布的卷积与 Fourier 变换

1. 经典结果

定义 5.3.3(**磨光核与磨光算子**) 设 $\alpha\in C_C^\infty(\mathbb{R}^n)$,$\alpha\geqslant 0$,$\int_{\mathbb{R}^n}\alpha(x)\mathrm{d}x=1$,且其支集等于半径为 1、球心在原点 0 的单位闭球 $\mathrm{supp}\alpha=\overline{B_1(0)}$. 令

$$\alpha_\varepsilon(x)=\varepsilon^{-n}\alpha\left(\frac{x}{\varepsilon}\right),\quad \varepsilon>0,$$

称 $\alpha_\varepsilon(x)$为一个**磨光核**. 显然,磨光核有如下性质:

$$\alpha_\varepsilon\in C_C^\infty(\mathbb{R}^n),\quad \alpha_\varepsilon\geqslant 0,\quad \int_{\mathbb{R}^n}\alpha_\varepsilon(x)\mathrm{d}x=1,$$

$$\mathrm{supp}\alpha_\varepsilon=\overline{B_\varepsilon(0)},\quad \lim_{\varepsilon\to 0}\{\mathrm{supp}\alpha_\varepsilon(x)\}=\{0\}.$$

取 $\varepsilon=\dfrac{1}{j}(j=1,2,\cdots)$,称函数序列

$$\alpha_j(x)=j^n\alpha(jx),\quad j=1,2,\cdots$$

为**正则化序列**,并称卷积

$$J_\varepsilon f(x)=f*\alpha_\varepsilon(x)=\int_{\mathbb{R}^n}f(t)\alpha_\varepsilon(x-t)\mathrm{d}t$$

为函数 $f(x)$的**磨光算子**. 磨光算子又称为 Fredrish 软化子.

定理 5.3.8(**磨光算子的性质**) 对于 $J_\varepsilon f(x)=f*\alpha_\varepsilon(x)$,有

(1) 若 $f\in L_{\mathrm{loc}}(\mathbb{R}^n)$,则 $f*\alpha_\varepsilon\in C^\infty(\mathbb{R}^n)$;

(2) 若 $f\in L_{\mathrm{loc}}(\mathbb{R}^n)$,且$\mathrm{supp}f=K\subset\mathbb{R}^n$ 为紧致集,则$\mathrm{supp}f*\alpha_\varepsilon=K_\varepsilon=\bigcup_{x\in K}B_\varepsilon(x)$;

(3) 若 $f\in C(\mathbb{R}^n)$,则$\lim\limits_{\varepsilon\to 0}f*\alpha_\varepsilon(x)=f(x)$在任一紧致集 $L\subset\mathbb{R}^n$ 上一致成立;

(4) 若 $f\in L^p(\mathbb{R}^n)$,$1\leqslant p<+\infty$,则 $\lim\limits_{\varepsilon\to 0}f*\alpha_\varepsilon\overset{L^p}{=}f$,亦即

$$\lim_{\varepsilon\to 0}\|f*\alpha_\varepsilon-f\|_{L^p(\mathbb{R}^n)}=0;$$

(5) 若 $f\in C_C^\infty(\mathbb{R}^n)$,则 $f*\alpha_\varepsilon\xrightarrow{D^*(\mathbb{R}^n)}f(\varepsilon\to 0)$,亦即,在每个有界集 $B\subset C_C^\infty(\mathbb{R}^n)$上

$$\lim_{\varepsilon\to 0}f*\alpha_\varepsilon=f$$

一致成立.

从定理可看出"磨光"算子的意义,也看出卷积在函数研究中的作用.

例 5.3.7 定理 5.3.8(4)中极限 $\lim\limits_{\varepsilon\to 0}\|f*\alpha_\varepsilon-f\|_{L^p(\mathbb{R}^n)}=0$ 与(5)中极限 $\lim\limits_{\varepsilon\to 0}f*\alpha_\varepsilon=f$ 的意义.

定理 5.3.8(4)中极限表示在 $L^p(\mathbb{R}^n)(1\leqslant p<+\infty)$ 中的平均极限,$f*\alpha_\varepsilon$ 相当于一个算子,当 $\varepsilon\to 0$ 时,此算子的极限是一个单位算子 I,即 $If=f$. (5)中的极限表示 $f*\alpha_\varepsilon$ 相当于一个算子,当 $\varepsilon\to 0$ 时,在每个有界集 $B\subset C_C^\infty(\mathbb{R}^n)$ 上,$f*\alpha_\varepsilon$ 以 f 为极限,而算子 $f*\alpha_\varepsilon$ 的极限是一个单位算子 I. 因此,卷积算子的极限相当于一个单位算子.

也要提醒的是,所考虑的卷积 $J_\varepsilon f(x)=f*\alpha_\varepsilon(x)$,其中的函数 $\alpha_\varepsilon(x)$ 具有紧致支集 $\mathrm{supp}\alpha_\varepsilon=\overline{B_\varepsilon(0)}$.

例 5.3.8 设一个线性系统的脉冲响应为 $h(t)$,进入系统的输入信号为 $f(t)$,试问输出信号是什么?

如图 5.3.1 所示,由物理知识可知,输入信号 $f(t)$ 在进入系统时有时滞 $t-\tau$,因此经过系统时,经脉冲响应混频,得到的输出信号为

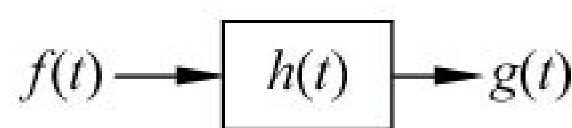


图 5.3.1

$$g(t)=f*h(t)=\int f(t-\tau)h(\tau)\mathrm{d}\tau,$$

因此卷积在信号分析中起重要作用.

2. 分布的卷积

对于函数 $f,g\in D(\mathbb{R})\subset D^*(\mathbb{R})$ 的卷积,亦即对于具有紧致支集的两个函数 f,g,它们的卷积

$$f*g(x)=\int_{\mathbb{R}}f(x-t)g(t)\mathrm{d}t=\int_{\mathbb{R}}f(y)g(x-y)\mathrm{d}y=\langle f(y),g(x-y)\rangle$$

可视为卷积算子 $f*g$ 对 $\varphi\in D(\mathbb{R})$ 的作用 $\langle f*g,\varphi\rangle$,亦即

$$\begin{aligned}\langle f*g,\varphi\rangle&=\left\langle\int_{\mathbb{R}}f(t)g(x-t)\mathrm{d}t,\varphi\right\rangle=\int_{\mathbb{R}}\varphi(x)\left(\int_{\mathbb{R}}f(t)g(x-t)\mathrm{d}t\right)\mathrm{d}x\\&=\int_{\mathbb{R}}f(t)\left(\int_{\mathbb{R}}\varphi(t+y)g(y)\mathrm{d}y\right)\mathrm{d}t\\&=\int_{\mathbb{R}}f(x)\left(\int_{\mathbb{R}}\varphi(x+y)g(y)\mathrm{d}y\right)\mathrm{d}x\\&=\langle f(x),\langle g(y),\varphi(x+y)\rangle\rangle,\end{aligned}$$

于是,启发我们定义分布的卷积为一个分布,满足 $\langle f*g,\varphi\rangle=\langle f(x),\langle g(y),\varphi(x+y)\rangle\rangle$.

定义 5.3.4(分布的卷积) 分为两种情形加以定义.

(1) 分布 $T\in D^*(\mathbb{R}^n)$ 与函数 $f\in D(\mathbb{R}^n)$ 的卷积

对于 $T\in D^*(\mathbb{R}^n)$,$f\in D(\mathbb{R}^n)$,定义卷积 $T*f$ 为一个分布,满足

$$\langle T * f, \varphi \rangle = \langle T_x, \langle f(y), \varphi(x+y) \rangle \rangle, \quad \forall \varphi \in D(\mathbb{R}^n).$$

也记为 $\langle T * f, \varphi \rangle = \langle T_x, \langle f_y, \varphi_{x+y} \rangle \rangle$, $\forall \varphi \in D(\mathbb{R}^n)$.

注 这样的定义有意义,因为 $f \in D(\mathbb{R}^n)$, $\forall \varphi \in D(\mathbb{R}^n)$,有

$$\psi(x) \equiv \langle f(y), \varphi(x+y) \rangle = \int_{\mathbb{R}^n} f(y)\varphi(x+y)\,\mathrm{d}y \in C^\infty(\mathbb{R}^n);$$

并且可化为

$$\psi(x) = \int_{\mathbb{R}^n} f(-y)\varphi(x-y)\,\mathrm{d}y = \int_{\mathbb{R}^n} \tilde{f}(y)\varphi(x-y)\,\mathrm{d}y = \tilde{f} * \varphi(x), \quad (5.3.1)$$

其中 $\tilde{f}(x) = f(-x)$. 又 $\mathrm{supp}\psi \subset \mathrm{supp}\,\tilde{f} + \mathrm{supp}\varphi$ 是一个紧致集,故 $\psi \in D(\mathbb{R}^n)$(因 $\varphi, \tilde{f} \in D(\mathbb{R}^n)$).

$T * f$ 是 $D(\mathbb{R}^n)$ 上的连续线性泛函,因为,$\langle T * f, \varphi \rangle = \langle T_x, \langle f(y), \varphi(x+y) \rangle \rangle$ 是线性的;并且当 $\varphi_j \in D(\mathbb{R}^n)$, $\varphi_j \xrightarrow{D(\mathbb{R}^n)} 0$ 时,有 $\langle f(y), \varphi_j(x+y) \rangle \xrightarrow{\mathbb{R}} 0$,从而有

$$\langle T * f, \varphi_j \rangle = \langle T_x, \langle f(y), \varphi_j(x+y) \rangle \rangle \xrightarrow{\mathbb{R}} 0.$$

这说明定义 $\langle T * f, \varphi \rangle = \langle T_x, \langle f_y, \varphi(x+y) \rangle \rangle$, $\forall \varphi \in D(\mathbb{R}^n)$,有意义.

(2) 分布 $T \in D^*(\mathbb{R}^n)$ 与分布 $S \in E^*(\mathbb{R}^n)$ 的卷积

设 $T \in D^*(\mathbb{R}^n)$, $S \in E^*(\mathbb{R}^n)$,即 $\mathrm{supp}S$ 为紧致集,则卷积 $T * S$ 定义为一个分布,满足

$$\langle T * S, \varphi \rangle = \langle T_x \otimes S_y, \varphi(x+y) \rangle, \quad \forall \varphi \in D(\mathbb{R}^n),$$

其中 $\langle T_x \otimes S_y, u(x,y) \rangle = \langle T_x, \langle S_y, u(x,y) \rangle \rangle = \langle S_y, \langle T_x, u(x,y) \rangle \rangle$, $\forall u \in D(\mathbb{R}^n \times \mathbb{R}^n)$,且 $\langle T \otimes S, \varphi \otimes \psi \rangle = \langle T, \varphi \rangle \langle S, \psi \rangle$ 是 T 与 S 的张量积.

这里只定义两个分布中有一个为紧致支集的卷积,两个分布都不具有紧致支集的情形已超出本课程的范围.

定理 5.3.9 (1) 当 $T \in D^*(\mathbb{R}^n)$ 时,若 $f \in D(\mathbb{R}^n)$,则卷积 $T * f(x)$ 是一个函数

$$T * f(x) = \langle T_y, f(x-y) \rangle, \quad x \in \mathbb{R}^n, \quad (5.3.2)$$

并且满足

① $T * f \in C^\infty(\mathbb{R}^n)$;

② $\mathrm{supp}(T * f) \subset \mathrm{supp}T + \mathrm{supp}f$;

(2) 当 $T \in D^*(\mathbb{R}^n)$ 时,若 $S \in E^*(\mathbb{R}^n)$,则卷积 $T * S$ 成为一个分布

$$\langle T * S, \varphi \rangle = \langle T_x, \langle S_y, \varphi(x+y) \rangle \rangle, \quad \forall \varphi \in D(\mathbb{R}^n), \quad (5.3.3)$$

并且满足

① $T * S \in D^*(\mathbb{R}^n)$;

② $\mathrm{supp}(T * S) \subset \mathrm{supp}T + \mathrm{supp}S$;

(3) 当 $T \in \mathbb{S}^*(\mathbb{R}^n)$ 时,若 $\varphi \in \mathbb{S}(\mathbb{R}^n)$,则卷积 $T * \varphi$ 成为一个函数

$$T * \varphi(x) = \langle T_y, \varphi(x-y) \rangle, \quad x \in \mathbb{R}^n, \quad (5.3.4)$$

并且满足

① $T*\varphi\in\mathbb{O}_M^n(\mathbb{R}^n)$，其中$\mathbb{O}_M^n(\mathbb{R}^n)\subset\mathbb{S}^*(\mathbb{R}^n)$是函数空间

$\mathbb{O}_M^n(\mathbb{R}^n)=\{f\in C^\infty(\mathbb{R}^n):\forall\alpha,\exists c_\alpha>0,\exists N_\alpha,\text{s. t. }|D^\alpha f(x)|\leqslant c_\alpha(1+|x|^2)^{N_\alpha}\}$；

② $(T*\varphi)^?=T^?\varphi^?$，且 $\operatorname{supp}(T*\varphi)\subset\operatorname{supp}T+\operatorname{supp}\varphi$；

(4) 当 $T\in\mathbb{S}^*(\mathbb{R}^n)$时，若 $S\in E^*(\mathbb{R}^n)$，则满足

① $S^?(\xi)=\langle S_x,\mathrm{e}^{-\mathrm{i}x\cdot\xi}\rangle,\xi\in\mathbb{R}^n$； $T*S\in\mathbb{S}^*(\mathbb{R}^n)$；

② $(T*S)^?=T^?S^?$，且 $\operatorname{supp}(T*S)\subset\operatorname{supp}T+\operatorname{supp}S$.

证 我们只证明(1)、(4)，其余情形可参看[5].

对于(1) $\forall T\in D^*(\mathbb{R}^n),f\in D(\mathbb{R}^n)$，由定义 5.3.4(1)，有

$$\langle T*f,\varphi\rangle=\langle T_x,\langle f(y),\varphi(x+y)\rangle\rangle,\quad\forall\varphi\in D(\mathbb{R}^n);$$

在该定义后面的注中已经阐明它有意义. 由于 $f\in D(\mathbb{R}^n),\varphi\in D(\mathbb{R}^n)$，所以

$$\langle f(y),\varphi(x+y)\rangle=\int_{\mathbb{R}^n}f(y)\varphi(x+y)\mathrm{d}y=\int_{\mathbb{R}^n}f(t-x)\varphi(t)\mathrm{d}t;$$

故

$$\langle T*f,\varphi\rangle=\langle T_x,\langle f(y),\varphi(x+y)\rangle\rangle=\left\langle T_x,\int_{\mathbb{R}^n}f(t-x)\varphi(t)\mathrm{d}t\right\rangle$$

$$=\left\langle\int_{\mathbb{R}^n}\langle T_x,f(t-x)\rangle\varphi(t)\mathrm{d}t\right\rangle=\langle\langle T_x,f(t-x)\rangle,\varphi(t)\rangle,$$

这里使用了“分布的作用与积分号交换次序”，其合理性是因为：可以将积分化为求和的极限，然后利用紧致集上收敛的一致性而交换两种运算的次序. 这样，在上式中，将 x 记为 y，将 t 记为 x，得到 $T*f(x)=\langle T_y,f(x-y)\rangle(x\in\mathbb{R}^n)$. 此即(5.3.2)式.

为证 $T*f\in C^\infty(\mathbb{R}^n)$，对于任意指标 $\alpha,\beta\in\mathbb{N}^n$，知

$$f\in D(\mathbb{R}^n)\Rightarrow\partial_x^\alpha\partial_y^\beta f(x-y)\in D(\mathbb{R}^n).$$

于是，例如当 $n=1$ 时，对于 $T\in D^*(\mathbb{R})$，任给 $h\in\mathbb{R},h\to0$，由

$$\frac{\mathrm{d}}{\mathrm{d}x}[T*f(x)]=\lim_{h\to0}\frac{1}{h}[T*f(x+h)-T*f(x)]$$

$$=\lim_{h\to0}\frac{1}{h}\{\langle T_y,f(x+h-y)\rangle-\langle T_y,f(x-y)\rangle\}\quad(\text{在 }D^*(\mathbb{R})\text{ 中})$$

$$=\lim_{h\to0}\left\langle T_y,\frac{f(x-y+h)-f(x-y)}{h}\right\rangle=\langle T_y,\partial_x f(x-y)\rangle.$$

对于 $p\in\mathbb{N}^n$，可类似求得 $D^p[T*f(x)]$. 从而 $T*f\in C^\infty(\mathbb{R}^n)$，此即(1)中的①.

为证(1)中的②，若 $x\notin\operatorname{supp}T+\operatorname{supp}f$，则当 $y\in\operatorname{supp}f$ 时，$x-y\notin\operatorname{supp}f$(否则，$x\in\{y\}+\operatorname{supp}f\subset\operatorname{supp}T+\operatorname{supp}f$)，亦即，$x\notin\operatorname{supp}T+\operatorname{supp}f$ 等价于$\operatorname{supp}T\cap\operatorname{supp}f(x-\cdot)=\varnothing$. 从而，$T*f$ 在 x 附近为 0. 这蕴含 $x\notin\operatorname{supp}(T*f)$，这就证明了 $\operatorname{supp}(T*f)\subset\operatorname{supp}T+\operatorname{supp}f$，即(1)中的②.

对于(4)中的①，为证第一个式子，由 $S\in E^*(\mathbb{R}^n)$，取 $\forall\varphi\in\mathbb{S}(\mathbb{R}^n)$，根据定义，有

$$\langle S^{?},\varphi\rangle=\langle S,\varphi^{?}\rangle=\langle S_x,\langle\varphi_\xi,\mathrm{e}^{-\mathrm{i}x\cdot\xi}\rangle\rangle=\langle S_x\otimes\varphi_\xi,\mathrm{e}^{-\mathrm{i}x\cdot\xi}\rangle=\langle\varphi_\xi\otimes S_x,\mathrm{e}^{-\mathrm{i}x\cdot\xi}\rangle$$
$$=\langle\varphi_\xi,\langle S_x,\mathrm{e}^{-\mathrm{i}x\cdot\xi}\rangle\rangle=\langle\langle S_x,\mathrm{e}^{-\mathrm{i}x\cdot\xi}\rangle,\varphi_\xi\rangle,$$

从而,$S^{?}(\xi)=\langle S_x,\mathrm{e}^{-\mathrm{i}x\cdot\xi}\rangle$得证.

对于①的第二个式子,当 $T\in\mathbb{S}^*(\mathbb{R}^n),S\in E^*(\mathbb{R}^n)$时,因$\mathbb{S}^*(\mathbb{R}^n)\subset D^*(\mathbb{R}^n)$,故由定义 5.3.4(2),$\langle T*S,\varphi\rangle=\langle T_x,\langle S_y,\varphi_{x+y}\rangle\rangle$. 再由 $S\in E^*(\mathbb{R}^n)$的构造定理 5.3.6(2),有

$$S\in E^*(\mathbb{R}^n)\Leftrightarrow S=\sum_{|p|\leqslant r}\partial^p f,\quad f\in C_C^r(\mathbb{R}^n),$$

f 是具有紧致支集的 r 次连续可导函数,且满足 $\mathrm{supp}f\subset(\mathrm{supp}S)_\varepsilon$. 令 $S_y=g(y)=\partial^p f(y)$,则

$$\psi(x)\equiv\langle S_y,\varphi(x+y)\rangle=\langle g(y),\varphi(x+y)\rangle=(-1)^{|p|}\int f(y)\partial_x^p\varphi(x+y)\mathrm{d}y.$$

利用积分号下求极限与一个称为 Peetre 的不等式[5],得到

$$|(1+|x|^2)^{\frac{k}{2}}\partial^q\psi(x)|\leqslant C\int|f(y)(1+|y|^2)^{\frac{k}{2}}||(1+|x+y|^2)^{\frac{k}{2}}\partial_x^{p+q}\varphi(x+y)|\mathrm{d}y$$
$$\leqslant C\sup_{\mathbb{R}^n}|(1+|x|^2)^{\frac{k}{2}}\partial_x^{p+q}\varphi(x)|,$$

由此推出,$\varphi\in\mathbb{S}(\mathbb{R}^n)$蕴含 $\psi(x)\in\mathbb{S}(\mathbb{R}^n)$,故 $\varphi\in\mathbb{S}(\mathbb{R}^n)$确定 $\psi(x)\equiv\langle g(y),\varphi(x+y)\rangle\in\mathbb{S}(\mathbb{R}^n)$仍为一个$\mathbb{S}^*(\mathbb{R}^n)$分布,从而,$\psi(x)\equiv\langle S_y,\varphi(x+y)\rangle\in\mathbb{S}(\mathbb{R}^n)$给出:当 $T\in\mathbb{S}^*(\mathbb{R}^n)$时,$\langle T_x,\langle S_y,\varphi(x+y)\rangle\rangle=\langle T_x*S_y,\varphi(x+y)\rangle$有意义,且 $T*S\in\mathbb{S}^*(\mathbb{R}^n)$.

为证(4)中的②,先计算$(T*S)^{?}$. 取 $\varphi\in D(\mathbb{R}^n)\subset\mathbb{S}(\mathbb{R}^n)$,则由 $T*S\in\mathbb{S}^*(\mathbb{R}^n)\subset D^*(\mathbb{R}^n)$与(5.3.3)式,有

$$\langle(T*S)^{?},\varphi\rangle=\langle T*S,\varphi^{?}\rangle=\langle T_x,\langle S_y,\varphi^{?}(x+y)\rangle\rangle.\tag{5.3.5}$$

为计算$\langle S_y,\varphi^{?}(x+y)\rangle$,由(4)的①,

$$\int_{\mathbb{R}^n}\mathrm{e}^{-\mathrm{i}x\cdot\xi}[(S_\xi^{?})(\xi)\cdot\varphi(\xi)]\mathrm{d}\xi=\int_{\mathbb{R}^n}\langle S_y,\mathrm{e}^{-\mathrm{i}y\cdot\xi}\rangle\varphi(\xi)\mathrm{e}^{-\mathrm{i}x\cdot\xi}\mathrm{d}\xi=\int_{\mathbb{R}^n}\langle S_y,\mathrm{e}^{-\mathrm{i}(x+y)\cdot\xi}\rangle\varphi(\xi)\mathrm{d}\xi$$
$$=\langle S_y,\int_{\mathbb{R}^n}\mathrm{e}^{-\mathrm{i}(x+y)\cdot\xi}\varphi(\xi)\mathrm{d}\xi\rangle=\langle S_y,\varphi^{?}(x+y)\rangle$$

给出

$$\langle S_y,\varphi^{?}(x+y)\rangle=(S^{?}\cdot\varphi)^{?}\equiv(S^{?}\varphi)^{?};$$

代入(5.3.5)式,得

$$\langle(T*S)^{?},\varphi\rangle=\langle T*S,\varphi^{?}\rangle=\langle T_x,\langle S_y,\varphi^{?}(x+y)\rangle\rangle$$
$$=\langle T,(S^{?}\varphi)^{?}\rangle=\langle T^{?},S^{?}\varphi\rangle=\langle T^{?}S^{?},\varphi\rangle,$$

这就是分布卷积的 Fourier 变换公式.

支集的包含关系与(1)的证明类似.

注 定理 5.3.9 的结论(4)中,用到了两个分布的 Fourier 变换的乘积. 类似于 5.1.2 节定义函数与分布的乘积,我们定义:分布 $T,S\in\mathbb{S}^*(\mathbb{R}^n)$的乘积 $T\cdot S\equiv TS$ 为满足

$$\langle T\cdot S,\varphi\rangle=\langle T,S\cdot\varphi\rangle,\quad\forall\varphi\in\mathbb{S}(\mathbb{R}^n)$$

的分布,若上式右边的线性泛函有意义.(见习题 25)

3. 分布卷积的性质

定理 5.3.10 对于 $T\in D^*(\mathbb{R}^n)$，$f,g\in D(\mathbb{R}^n)$，$p\in\mathbb{N}^n$，$\delta\in E^*(\mathbb{R}^n)$，则

(1) $(T*f)*g=T*(f*g)$；$(T*\delta)*g=T*(\delta*g)$； （结合律）

(2) $T*f=f*T$； （交换律）

(3) $T*\delta=\delta*T=T$； （单位元）

(4) $\tau_hT=(\tau_h\delta)*T$； （分布的平移）

(5) $\tau_h(T*f)=(\tau_hT)*f=T*(\tau_hf)$； （卷积的平移）

(6) $D^p(T*f)=D^pT*f=T*D^pf$. （卷积的导数）

证 对于(1)，由卷积的定义，有

$$\langle T*f,\varphi\rangle=\langle T_x,\langle f_y,\varphi_{x+y}\rangle\rangle,\quad \forall\varphi\in D(\mathbb{R}^n).$$

任取 $\varphi\in D(\mathbb{R}^n)$，则由(1)式左边出发，得

$$\begin{aligned}\langle(T*f)*g,\varphi\rangle&=\langle(T*f)_x,\langle g_y,\varphi_{x+y}\rangle\rangle=\langle T_x,\langle f_y,\langle g_y,\varphi_{x+y}\rangle\rangle\rangle\\&=\langle T_x,\langle(f*g)_y,\varphi_{x+y}\rangle\rangle=\langle T_x*(f*g)_y,\varphi_{x+y}\rangle\\&=\langle T*(f*g),\varphi\rangle;\end{aligned}$$

$(T*\delta)*g=T*(\delta*g)$可类似得到. 从而(1)得证.

为证(2)，分为两步.

第一步，设 $T_1,T_2\in D^*(\mathbb{R}^n)$，若 $T_1*\varphi=T_2*\varphi$，$\forall\varphi\in D(\mathbb{R}^n)$，则 $T_1=T_2$.

事实上，$\forall\varphi\in D(\mathbb{R}^n)$，由定理 5.3.9 得

$$T*f(x)=\langle T_y,f(x-y)\rangle=\langle T_y,\tilde f(y-x)\rangle,\quad x\in\mathbb{R}^n,$$

在上式中令 $x=0$，得 $T*f(0)=\langle T_y,\tilde f(y)\rangle$；也有 $T*\tilde f(0)=\langle T_y,f(y)\rangle=\langle T,f\rangle$. 故

$$\langle T_1,\varphi\rangle=(T_1*\tilde\varphi)(0)=(T_2*\tilde\varphi)(0)=\langle T_2,\varphi\rangle,\quad\forall\varphi\in D(\mathbb{R}^n).$$

从而有 $T_1=T_2$.

第二步，设 $T\in D^*(\mathbb{R}^n)$，$f\in D(\mathbb{R}^n)$，则 $T*f=f*T$.

事实上，任取 $\varphi,\psi\in D(\mathbb{R}^n)$，则 $\varphi*\psi=\psi*\varphi\in D(\mathbb{R}^n)$. 于是

$$(T*f)*(\varphi*\psi)\underset{D(\mathbb{R}^n)\text{ 中的交换律}}{=}(T*f)*(\psi*\varphi)\underset{\text{本定理中的结合律(1)}}{=}T*(f*\psi*\varphi)$$

$$\begin{aligned}&=T*(f*\psi)*\varphi=T*\varphi*(f*\psi)\\&=(T*\varphi)*(f*\psi)=(f*\psi)*(T*\varphi)\\&=f*\psi*T*\varphi\underset{\psi*T=T*\psi}{=}f*T*\psi*\varphi\underset{\psi*\varphi=\varphi*\psi}{=}(f*T)*(\varphi*\psi)\end{aligned}$$

继而，由第一步得(2).

为证(3),即 $T*\delta=\delta*T=T$,也分两步.

第一步,当 $T=f\in D(\mathbb{R}^n)$时,证 $\delta*f=f$.

由定义

$$\delta*f(x)=\langle\delta,f(y+x)\rangle=f(x+y)\big|_{y=0}=f(x),$$

又将 δ 视为 $\delta\in D^*(\mathbb{R}^n)$,而 $f\in D(\mathbb{R}^n)$,则 $f=\delta*f=f*\delta$.

第二步,当 $T\in D^*(\mathbb{R}^n)$,由 $(T*\delta)*\varphi=T*(\delta*\varphi)=T*\varphi$,得 $T*\delta=\delta*T=T$.

此性质使我们可以将 $\delta\in D^*(\mathbb{R}^n)$视为 $D^*(\mathbb{R}^n)$中卷积运算 $*$ 的单位元.

为证(4),即 $\tau_hT=(\tau_h\delta)*T$,任取 $\varphi\in D(\mathbb{R}^n)$,则

$$\begin{aligned}\langle(\tau_h\delta)*T,\varphi\rangle&=\langle T*(\tau_h\delta),\varphi\rangle=\langle T_x,\langle(\tau_h\delta)_y,\varphi_{x+y}\rangle\rangle=\langle T_x,(\tau_h\delta)_y*\varphi_{x+y}\rangle\\&=\langle T_x,\varphi_{x+h}\rangle=\langle T,\tau_{-h}\varphi\rangle=\langle\tau_hT,\varphi\rangle.\end{aligned}$$

关于(5),为证 $\tau_h(T*f)=(\tau_hT)*f=T*(\tau_hf)$,任取 $\varphi\in D(\mathbb{R}^n)$,则

$$\begin{aligned}\langle\tau_h(T*f),\varphi\rangle&=\langle T*f,\tau_{-h}\varphi\rangle=\langle(T*f)_x,\varphi(x+h)\rangle\\&=\langle T_x,\langle f_y,\varphi_{x+y+h}\rangle\rangle=\langle T_x,\langle f_y,\tau_{-h}\varphi_{x+y}\rangle\rangle=\langle T_x,\langle\tau_hf_y,\varphi_{x+y}\rangle\rangle\\&=\langle T_x*\tau_hf_y,\varphi_{x+y}\rangle=\langle T*\tau_hf,\varphi\rangle,\end{aligned}$$

进而,由交换律 $\tau_h(T*f)=\tau_h(f*T)$与同样的推理,得

$$\tau_h(f*T)=(\tau_hf)*T=T*(\tau_hf),$$

从而(5)得证.

性质(5)表明卷积对于平移具有不变性.

最后,证明(6),$D^p(T*f)=D^pT*f=T*D^pf$. 任取 $\varphi\in D(\mathbb{R}^n)$,则对于任取的 $p=\alpha\in\mathbb{N}^n$,

$$\begin{aligned}\langle\partial^\alpha(T*f),\varphi\rangle&=(-1)^{|\alpha|}\langle T*f,\partial^\alpha\varphi\rangle=\langle T_x,(-1)^{|\alpha|}\langle f_y,\partial_y^\alpha\varphi_{x+y}\rangle\rangle\\&=\langle T_x,\langle\partial_y^\alpha f_y,\varphi_{x+y}\rangle\rangle=\langle T*\partial^\alpha f,\varphi\rangle;\end{aligned}$$

此即 $D^\alpha(T*f)=T*D^\alpha f$; 类似地,由交换律 $D^\alpha(T*f)=D^\alpha(f*T)$,推理得

$$D^\alpha(T*f)=f*D^\alpha T=D^\alpha T*f,$$

从而,定理全部得证.

分布卷积有如下的重要公式.

定理 5.3.11 对于 $T\in D^*(\mathbb{R}^n)$,$\varphi\in D(\mathbb{R}^n)$,有$\langle T,\varphi\rangle=(T_x*\widetilde{\varphi})(0)$.

4. $E^*(\mathbb{R}^n)$上的赋半范代数

定理 5.3.12 (1) 对于 $T*\alpha_\varepsilon(x)$,若 $T\in D^*(\mathbb{R}^n)$,则在 $D^*(\mathbb{R}^n)$的强拓扑下成立

$$\lim_{\varepsilon\to0}T*\alpha_\varepsilon=T.$$

(2) 分布空间$(E^*(\mathbb{R}^n),+,\alpha\cdot,\{p_{m,j}\},*)$是一个具有单位元 $\delta\in E^*(\mathbb{R}^n)$的可结合、可交换的赋半范代数,且 $\delta*T=T*\delta$, $\forall\, T\in E^*(\mathbb{R}^n)$.

5.4 小波分析

5.4.1 小波变换的引入

Fourier 变换在科学技术领域中起着极其重要的作用，持续了近 200 年. 然而，它也有不能满足科技需求的地方，为此，数学家改进了 Fourier 变换的不足之处. 1984 年，由美国数学家 A. Grossmann 与 J. Morlet 引入了小波(wavelets)变换. 此后，各国数学家纷纷跟上，奠基性工作、研究性成果、有效的应用大量涌现，形成了小波研究的热潮，至今小波分析仍然处于基础数学与应用数学研究的前沿. 本节以介绍一维小波理论为例.

1. Gabor 变换

在 5.2 节和 5.3 节中，我们分别定义了 $f\in L^p(\mathbb{R})(1\leqslant p\leqslant 2)$，与 Schwartz 分布的 Fourier 变换，并研究了它们的性质. 从 $f\in L^1(\mathbb{R})$的 Fourier 变换的定义

$$f^?(\xi)=\int_{\mathbb{R}}f(t)\mathrm{e}^{-\mathrm{i}t\xi}\mathrm{d}t,\quad \xi\in\mathbb{R}$$

可以看到，为了由 Fourier 变换研究一个信号 $f(t)\in L^1(\mathbb{R})$的谱性质，必须获得该信号在时域$\mathbb{R}=(-\infty,+\infty)$中的全部信息，否则无法计算其频谱 $f^?(\xi)$. 另一方面，信号 $f(t)$即使在一个小邻域$(t_0-\Delta t,t_0+\Delta t)$中有一些变化，也会影响到整个频谱的值. 例如，对于支撑在 t_0 的分布 $\delta_{t_0}=\tau_{t_0}\delta$，由定理 5.3.3(1)，有$(\delta_{t_0})^?=\mathrm{e}^{-\mathrm{i}t_0\omega}$，$\omega\in\mathbb{R}$，这说明支撑在 t_0 的分布 $\delta_{t_0}=\tau_{t_0}\delta$ 的 Fourier 变换竟然能覆盖整个频域.

为了提取信号 $f(t)$的 Fourier 变换 $f^?(\xi)$的局部信息，希望不需要整个频域的值，1946 年，B. Gabor，引入了“时域局部化”的 Gabor 变换

$$G_b^a f(\xi)=\int_{\mathbb{R}}f(t)\mathrm{e}^{-\mathrm{i}\xi t}g_a(t-b)\mathrm{d}t,$$

这里 g_a 是 Gabor 引进的“权函数”，$g_a(t)=\dfrac{1}{2\sqrt{\pi a}}\mathrm{e}^{-\frac{t^2}{4a}}$，$a>0$.

由于权函数的引进，使得 $f^?$ 被局部化了，这是因为由 Gauss 函数 e^{-x^2} 的性质以及其 Fourier 变换的性质所引起的. Gauss 函数的 Fourier 变换为

$$\int_{\mathbb{R}}\mathrm{e}^{-x^2}\mathrm{e}^{-\mathrm{i}x\xi}\mathrm{d}x=\sqrt{\frac{\pi}{a}}\mathrm{e}^{-\frac{\xi^2}{4a}},\quad \xi\in\mathbb{R},$$

从而得到

$$\int_{\mathbb{R}}g_a(t-b)\mathrm{d}b=\frac{1}{2\sqrt{\pi a}}\int_{\mathbb{R}}\exp\left\{-\frac{(t-b)^2}{4a}\right\}\mathrm{d}b=\frac{1}{\sqrt{\pi}}\int_{\mathbb{R}}\mathrm{e}^{-\xi^2}\mathrm{d}\xi=1,$$

于是

$$\int_{\mathbb{R}}(G_b^a f)(\xi)\mathrm{d}b=\int_{\mathbb{R}}\left(\int_{\mathbb{R}}f(t)\mathrm{e}^{-\mathrm{i}\xi t}g_a(t-b)\mathrm{d}t\right)\mathrm{d}b$$
$$=\int_{\mathbb{R}}f(t)\mathrm{e}^{-\mathrm{i}\xi t}\left(\int_{\mathbb{R}}g_a(t-b)\mathrm{d}b\right)\mathrm{d}t=\int_{\mathbb{R}}f(t)\mathrm{e}^{-\mathrm{i}\xi t}\cdot 1\mathrm{d}t$$
$$=\int_{\mathbb{R}}f(t)\mathrm{e}^{-\mathrm{i}\xi t}\mathrm{d}t=f^?(\xi),\quad \xi\in\mathbb{R}.$$

这表明 $f^?(\xi)=\int_{\mathbb{R}}G_a^b f(\xi)\mathrm{d}b$ 与 Gabor 变换 $G_b^a f(\xi)$ 中的参数 a 无关，而是在 $t=b$ 被"局部化"了，可以把 $f^?(\xi)$ 的值看做 $g_a(t-b)$ 集中在 $t=b$ 附近，因此提取了信号的 Fourier 变换(频谱)的局部信息，也就是 Gabor 变换集 $\{G_b^a f:b\in\mathbb{R}\}$ 分解了 $f^?(\xi)$，提取了 $f(t)$ 的局部谱信息.

令 $G_{b,\xi}^a(t)=\mathrm{e}^{-\mathrm{i}\xi t}g_a(t-b)$，我们改写 $G_b^a f$ 为

$$(G_b^a f)(\xi)=\int_{\mathbb{R}}f(t)\mathrm{e}^{-\mathrm{i}\xi t}g_a(t-b)\mathrm{d}t$$
$$\equiv\int_{\mathbb{R}}f(t)G_{b,\xi}^a(t)\mathrm{d}t=\langle G_{b,\xi}^a,f\rangle$$

这就赋予了 $G_b^a f$ 以新的意义：对 $f(t)$ 开一个"窗"，用"窗函数"$G_{b,\xi}^a(t)$ 把 $f^?$ 限制在局部范围内.

新的问题对 Gabor 变换提出挑战：(1) 如何由 g_a 自身决定"窗"的大小？(2) 由于 Gauss 函数性质的限制，使得窗函数的灵活性受到限制，如何克服？因此，为解决这些问题，小波分析蓬勃崛起.

2. 窗函数

以下使用物理学中常用的记号，t 为时间，ω 为信号 $g(t)$ 的频率，$g^?(\omega)$ 为信号的 Fourier 变换，即频谱. (t,ω) 为时-频空间，亦称相空间(phase space).

定义 5.4.1(窗函数) 若函数 $g(t)\in L^2(\mathbb{R})$，$tg(t)\in L^2(\mathbb{R})$，则称 $g(t)$ 为**窗函数(window function)**；当 $g(t)$ 与 $g^?(\omega)$ 都是窗函数时，称

$$(t_0,\omega_0)=\left(\frac{1}{\|g\|_2^2}\int_{\mathbb{R}}t|g(t)|^2\mathrm{d}t,\frac{1}{\|g^?\|_2^2}\int_{\mathbb{R}}\omega|g^?(\omega)|^2\mathrm{d}\omega\right)$$

为相空间中的**中心(center)**，其中 $\|g\|_2\equiv\|g\|_{L^2(\mathbb{R})}$. 并且分别称

$$\Delta_g=\left\{\frac{\int_{\mathbb{R}}(t-t_0)^2|g(t)|^2\mathrm{d}t}{\|g\|_2^2}\right\}^{\frac{1}{2}},\quad \Delta_{g^?}=\left\{\frac{\int_{\mathbb{R}}(\omega-\omega_0)^2|g(\omega)|^2\mathrm{d}\omega}{\|g^?\|_2^2}\right\}^{\frac{1}{2}}$$

为信号 $g(t)$ 的**时宽(time width)**与**频宽(frequency width)**；称相空间 (t,ω) 中以 (t_0,ω_0) 为中心、以 $2\Delta_g$ 与 $2\Delta_{g^?}$ 为边长的矩形

$$[T_g^0;\Omega_{g^?}^0]\equiv T_g^0\times\Omega_{g^?}^0=[t_0-\Delta_g,t_0+\Delta_g;\omega_0-\Delta_{g^?},\omega_0+\Delta_{g^?}]$$

为 $g(t)$ 的**时-频窗口(time-frequency window)**.

3. 窗函数的简单性质

命题 5.4.1(测不准原理) 设 $g\in L^2(\mathbb{R})$,若其自身 $g(t)$ 及其 Fourier 变换 $g^?(\omega)$ 都是窗函数,则

$$\Delta_g\cdot\Delta_{g^?}\geqslant\frac{1}{2};$$

等号成立的充分必要条件是

$$g(t)=c\mathrm{e}^{\mathrm{i}\alpha t}\frac{1}{2\sqrt{\pi a}}\exp\left\{-\frac{(t-b)^2}{4a}\right\},\quad c\neq 0,\ \alpha\in\mathbb{R};\ a>0,\ b\in\mathbb{R}.$$

注 由定义知,时-频窗口的面积为 $4\Delta_g\Delta_{g^?}$,于是,测不准原理告诉我们,任何窗函数的**时-频窗口的面积都大于或等于 2**;并且,**任何窗函数的时-频窗口的面积都不小于 Gauss 函数所决定的时-频窗口的面积.**

这个命题的数学证明并不困难,但是其物理意义却很重要.

命题 5.4.2(平移伸缩性质) 设 $g,g^?$ 是窗函数,则平移伸缩函数

$$g^{ba}(x)=\frac{1}{\sqrt{a}}g\left(\frac{x-b}{a}\right),\quad a>0,b\in\mathbb{R}$$

与其 Fourier 变换

$$(g^{ba}(\circ))^?(\omega)=\sqrt{|a|}\mathrm{e}^{-\mathrm{i}b\omega}g^?(a\omega),\quad a>0,b\in\mathbb{R},\omega\in\mathbb{R}$$

也是窗函数;这里 $g^{ba}(x)$ 的中心是 $b+at_0$,半径是 $a\Delta_g$;$(g^{ba})^?(\omega)$ 的中心是 $\dfrac{\omega_0}{a}$,半径是 $\dfrac{\Delta_{g^?}}{a}$,而时-频窗口是 $[T^0_{g^{ba}};\Omega^0_{(g^{ba})^?}]=\left[b+at_0-a\Delta_g,b+at_0+a\Delta_g;\dfrac{\omega_0}{a}-\dfrac{\Delta_{g^?}}{a},\dfrac{\omega_0}{a}+\dfrac{\Delta_{g^?}}{a}\right]$;时-频窗口的面积仍然是 $4\Delta_g\Delta_{g^?}$.

把信号限制在时间窗中,称为信号的**时间局部化(time localization)**;把信号的频谱限制在频率窗中,称为**信号的频率局部化(frequency localization)**.

5.4.2 连续小波变换

现在引入连续小波变换的概念,并研究其性质.

1. 连续小波变换

定义 5.4.2(容许条件) 若函数 $\psi\in L^2(\mathbb{R})$ 满足

$$0<c_\psi=\int_{\mathbb{R}}\frac{|\psi^?(\omega)|^2}{|\omega|}\mathrm{d}\omega<+\infty,$$

则称 ψ 满足**容许条件(admissible condition)**,并称 $\psi\in L^2(\mathbb{R})$ 为**一个基小波(base wavelet)**.

为了方便起见,不失一般性,假设 $\|\psi\|_2=1$.

对于任意一个基小波 $\psi\in L^2(\mathbb{R})$,任意的 $a,b\in\mathbb{R}$,$a\neq 0$,作平移伸缩,得到

$$\psi^{ba}(x)=\frac{1}{\sqrt{|a|}}\psi\left(\frac{x-b}{a}\right), \tag{5.4.1}$$

这是一个**双指标族**,称为由基小波 ψ 产生的**小波族(wavelet family)**;显然,有

$$(\psi^{ba})^{\wedge}(\omega)=\sqrt{|a|}\,\mathrm{e}^{-\mathrm{i}b\omega}\psi^{\wedge}(a\omega). \tag{5.4.2}$$

定义 5.4.3(连续小波变换) 对于任意函数 $f\in L^2(\mathbb{R})$,由小波族(5.4.1)生成的算子

$$(\mathrm{W}_\psi f)(b,a)=\frac{1}{\sqrt{|a|}}\int_{\mathbb{R}} f(t)\,\overline{\psi\left(\frac{t-b}{a}\right)}\mathrm{d}t\equiv\langle f,\psi^{ba}\rangle,\quad a,b\in\mathbb{R},a\neq 0 \tag{5.4.3}$$

称为 f 的**连续小波变换(continuous wavelet transform)**,简称**小波变换**.

定理 5.4.1(连续小波变换的等价表示) 连续小波变换(5.4.3)是一个卷积算子,即

$$(\mathrm{W}_\psi f)(b,a)=\langle f,\psi^{ba}\rangle=\frac{1}{\sqrt{|a|}}f*\overline{\psi(a^{-1}\circ)}(b),$$

其中 $f\in L^2(\mathbb{R})$,$a,b\in\mathbb{R}$,$a\neq 0$.

定理 5.4.2(小波族的时-频窗口) 对于 $a>0$,$b\in\mathbb{R}$,小波族(5.4.1)的时-频窗口为

$$T^0_{\psi^{ba}}\times\Omega^0_{(\psi^{ba})^{\wedge}}=\left[b+at_0-a\Delta_\psi,b+at_0+a\Delta_\psi;\frac{\omega_0}{a}-\frac{1}{a}\Delta_{\psi^{\wedge}},\frac{\omega_0}{a}+\frac{1}{a}\Delta_{\psi^{\wedge}}\right].$$

证 因为 ψ 的时-频窗口为

$$T^0_\psi\times\Omega^0_{\psi^{\wedge}}=[t_0-\Delta_\psi,t_0+\Delta_\psi;\omega_0-\Delta_{\psi^{\wedge}},\omega_0+\Delta_{\psi^{\wedge}}],$$

再由命题 5.4.2 与式(5.4.2),定理得证.

2. 容许条件的意义

定理 5.4.3 设 ψ 是一个基小波,则由它决定的连续小波变换(5.4.3)满足

$$J=\int_{\mathbb{R}}\int_{\mathbb{R}}(\mathrm{W}_\psi f)(b,a)\cdot\overline{(\mathrm{W}_\psi g)(b,a)}\frac{\mathrm{d}a}{a^2}\mathrm{d}b=c_\psi\langle f,g\rangle,$$

其中 $f,g\in L^2(\mathbb{R})$,c_ψ 由定义 5.4.2 给出.

证 由

$$I\equiv\int_{\mathbb{R}}(\mathrm{W}_\psi f)(b,a)\cdot\overline{(\mathrm{W}_\psi g)(b,a)}\mathrm{d}b \qquad \text{(定义)}$$

$$=\frac{1}{|a|}\int_{\mathbb{R}}\left\{\int_{\mathbb{R}}f(t)\,\overline{\psi\left(\frac{t-b}{a}\right)}\mathrm{d}t\cdot\int_{\mathbb{R}}\overline{g(s)}\psi\left(\frac{s-b}{a}\right)\mathrm{d}s\right\}\mathrm{d}b \qquad \text{(Plancherel 型公式)}$$

$$=\int_{\mathbb{R}}\left\{\frac{|a|^{\frac12}}{\sqrt{2\pi}}\int_{\mathbb{R}}f^{\wedge}(\omega)\mathrm{e}^{\mathrm{i}b\omega}\overline{\psi^{\wedge}(a\omega)}\mathrm{d}\omega\cdot\frac{|a|^{\frac12}}{\sqrt{2\pi}}\int_{\mathbb{R}}\overline{g^{\wedge}(\tau)}\mathrm{e}^{-\mathrm{i}b\tau}\psi^{\wedge}(a\tau)\mathrm{d}\tau\right\}\mathrm{d}b$$

$$=|a|\int_{\mathbb{R}}\left\{\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}F(\omega)\mathrm{e}^{\mathrm{i}b\omega}\mathrm{d}\omega\cdot\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\overline{G(\tau)}\mathrm{e}^{-\mathrm{i}b\tau}\mathrm{d}\tau\right\}\mathrm{d}b$$

$$\left(令\ F(\omega)=f^{?}(\omega)\overline{\psi^{?}(a\omega)}, G(\tau)=g^{?}(\tau)\overline{\psi^{?}(a\tau\right.$$

$$=|a|\int_{\mathbb{R}}\left\{\frac{1}{\sqrt{2\pi}}\overline{\int_{\mathbb{R}}\overline{F(\omega)}\mathrm{e}^{-\mathrm{i}b\omega}\mathrm{d}\omega}\cdot\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}\overline{G(\tau)}\mathrm{e}^{-\mathrm{i}b\tau}\mathrm{d}\tau\right\}\mathrm{d}b$$

$$=\frac{|a|}{2\pi}\int_{\mathbb{R}}\overline{\overline{F(\circ)}^{?}(b)}\,\overline{G(\circ)}^{?}(b)\mathrm{d}b=|a|\int_{\mathbb{R}}F(x)\overline{G(x)}\mathrm{d}x.\text{(Plancherel 型公式}$$

于是

$$\begin{aligned}J&=\int_{\mathbb{R}}I\frac{\mathrm{d}a}{a^2}=\int_{\mathbb{R}}\int_{\mathbb{R}}(W_{\psi}f)(b,a)\cdot\overline{(W_{\psi}g)(b,a)}\mathrm{d}b\cdot\frac{\mathrm{d}a}{a^2}\\&=\int_{\mathbb{R}}\frac{a^2}{|a|}\int_{\mathbb{R}}\overline{G(x)}F(x)\mathrm{d}x\cdot\frac{\mathrm{d}a}{a^2}\\&=\int_{\mathbb{R}}\frac{1}{|a|}\left\{\int_{\mathbb{R}}\overline{g^{?}(x)}\psi^{?}(ax)\cdot f^{?}(x)\overline{\psi^{?}(ax)}\mathrm{d}x\right\}\mathrm{d}a,\end{aligned}$$

再由 Fubini 定理得

$$J=\int_{\mathbb{R}}f^{?}(x)\overline{g^{?}(x)}\left\{\int_{\mathbb{R}}\frac{|\psi^{?}(ax)|^2}{|a|}\mathrm{d}a\right\}\mathrm{d}x=c_{\psi}\langle f^{?},g^{?}\rangle=c_{\psi}\langle f,g\rangle,$$

这就是要证明的结论.

这个结果也可写为

$$\langle W_{\psi}f,W_{\psi}g\rangle=c_{\psi}\langle f,g\rangle,$$

左边是二元函数$(W_{\psi}f)(b,a)$在“权元”$\frac{\mathrm{d}a}{a^2}\mathrm{d}b$之下的“内积”,由此可见,容许条件中的值 c_{ψ} 为 0 与$+\infty$时,小波变换将无意义.

3. 反演公式

定理 5.4.4(反演公式) 设 ψ 是一个基小波,则由它决定的连续小波变换(5.4.3)式满足反演公式

$$f(x)=\frac{1}{c_{\psi}}\int_{\mathbb{R}}\int_{\mathbb{R}}(W_{\psi}f)(b,a)\psi^{ba}(x)\frac{\mathrm{d}a}{a^2}\mathrm{d}b$$

对于 $f\in L^2(\mathbb{R})$的任意连续点 $x\in\mathbb{R}$ 都成立.

证 在定理 5.4.3 与(5.4.3)式中,取 g 为 Gauss 函数

$$g_a(t-b)=\frac{1}{2\sqrt{\pi a}}\mathrm{e}^{-\frac{(t-b)^2}{4a}},\quad a>0,b\in\mathbb{R},$$

得

$$\begin{aligned}I&=\frac{1}{c_{\psi}}\lim_{a\to0+}\int_{\mathbb{R}}\int_{\mathbb{R}}(W_{\psi}f)(b,a)\cdot\overline{(W_{\psi}g_a(\circ-b))}\frac{\mathrm{d}a}{a^2}\mathrm{d}b\\&=\frac{1}{c_{\psi}}\lim_{a\to0+}\int_{\mathbb{R}}\int_{\mathbb{R}}(W_{\psi}f)(b,a)\cdot\overline{\langle g_a(\circ-b),\psi^{ba}\rangle}\frac{\mathrm{d}a}{a^2}\mathrm{d}b\end{aligned}$$

$$= \frac{1}{c_\psi}\int_{\mathbb{R}}\int_{\mathbb{R}} (W_\psi f)(b,a) \cdot \lim_{a\to 0+} g_a * \overline{\psi^{ba}(x)} \frac{\mathrm{d}a}{a^2}\mathrm{d}b,$$

利用卷积的性质,知 $\lim\limits_{a\to 0+} g_a * \psi^{ba}(x) = \psi^{ba}(x)$,代入上式,得到

$$I = \frac{1}{c_\psi}\int_{\mathbb{R}}\int_{\mathbb{R}} (W_\psi f)(b,a) \cdot \psi^{ba}(x) \frac{\mathrm{d}a}{a^2}\mathrm{d}b.$$

再由定理5.4.3,得到

$$I = \frac{1}{c_\psi}\lim_{a\to 0+}\{c_\psi\langle f,g\rangle\} = \lim_{a\to 0+}\langle f, g_a(\circ - x)\rangle = \lim_{a\to 0+} g_a * f(x) = f(x),$$

于是,定理5.4.4得证.

定理5.4.5(基小波的判别与反演公式) 设 $\psi\in L^2(\mathbb{R})$,$\|\psi\|_2=1$,且满足

$$\int_{(0,+\infty)}\frac{|\hat{\psi}(\omega)|^2}{\omega}\mathrm{d}\omega = \int_{(0,+\infty)}\frac{|\hat{\psi}(-\omega)|^2}{|-\omega|}\mathrm{d}\omega = \frac{1}{2}c_\psi < +\infty,$$

则 ψ 是一个基小波,且由它决定的连续小波变换(5.4.3)的反演公式

$$f(x) = \frac{2}{c_\psi}\int_{(0,+\infty)}\left[\int_{\mathbb{R}}(W_\psi f)(b,a)\left\{\frac{1}{\sqrt{a}}\psi\left(\frac{x-b}{a}\right)\right\}\mathrm{d}b\right]\frac{\mathrm{d}a}{a^2},\quad a>0, b\in\mathbb{R}$$

对于 $f\in L^2(\mathbb{R})$ 成立.

下面的定理说明基小波的一个特性.

定理5.4.6 设 ψ 是一个基小波,亦即,满足 $\psi(t)\in L^2(\mathbb{R})$,且 $0<c_\psi<+\infty$;若 $\hat{\psi}$ 在点 $\omega=0$ 连续,则 $\int_{\mathbb{R}}\psi(t)\mathrm{d}t = 0$.

证 由

$$0 < c_\psi = \int_{\mathbb{R}}\frac{|\hat{\psi}(\omega)|^2}{|\omega|}\mathrm{d}\omega = \int_{(-\infty,0)}\frac{|\hat{\psi}(\omega)|^2}{|\omega|}\mathrm{d}\omega + \int_{(0,+\infty)}\frac{|\hat{\psi}(\omega)|^2}{|\omega|}\mathrm{d}\omega < +\infty,$$

得到 $\dfrac{|\hat{\psi}(\omega)|^2}{|\omega|}\in L^1(\mathbb{R}^+)$,从而 $\dfrac{|\hat{\psi}(\omega)|}{|\omega|^{\frac{1}{2}}}\in L^2(\mathbb{R}^+)$,这说明在 $\omega=0$ 附近 $\dfrac{|\hat{\psi}(\omega)|}{|\omega|^{\frac{1}{2}}}$ 是几乎处处有限的,因此,由 $\lim\limits_{|\omega|\to 0}\dfrac{|\hat{\psi}(\omega)|}{|\omega|^{\frac{1}{2}}}<+\infty$ 必定得到极限 $\lim\limits_{|\omega|\to 0}|\hat{\psi}(\omega)|=0$;再由 $\hat{\psi}(\omega)$ 在点 $\omega=0$ 的连续性假设,立即得到 $\hat{\psi}(0)=0$,亦即 $\int_{\mathbb{R}}\psi(t)\mathrm{d}t = 0$.

4. 基小波的例

例5.4.1 Mexico帽子 $\varphi(x)=\begin{cases}1-|x|, & |x|\leqslant 1,\\ 0, & |x|>1\end{cases}$ 是基小波.

例 5.4.2 Haar 函数 $\psi(t)=\begin{cases}1, & 0\leqslant t<\frac{1}{2},\\ -1, & \frac{1}{2}\leqslant t\leqslant 1,\\ 0, & t<\frac{1}{2}, t>1\end{cases}$ 是基小波.

例 5.4.3 B-样条函数 $\varphi(x)=\begin{cases}\frac{1}{2}(x+1)^2, & -1\leqslant x\leqslant 0,\\ \frac{3}{4}-\left(x-\frac{1}{2}\right)^2, & 0\leqslant x\leqslant 1,\\ \frac{1}{2}(x-1)^2, & 1\leqslant x\leqslant 2,\\ 0, & x<-1, x>2\end{cases}$ 是基小波.

5.4.3 离散小波变换

现在引入离散小波变换的概念,并研究其性质.

1. 直交小波

定义 5.4.4(直交基小波) 设函数 $\psi\in L^2(\mathbb{R})$, $\|\psi\|_2=1$. 令

$$\psi_{j,k}(x)=2^{\frac{j}{2}}\psi(2^j-k),\quad j,k\in\mathbb{Z},$$

若

(1) $\langle\psi_{j,k},\psi_{l,m}\rangle=\delta_{j,l}\delta_{k,m}$, $\delta_{j,l}=\begin{cases}1, & j=l,\\ 0, & j\neq l;\end{cases}$

(2) $\forall f\in L^2(\mathbb{R})$,在 $L^2(\mathbb{R})$ 范数收敛意义下成立

$$f(x)=\sum_{j,k=-\infty}^{+\infty}c_{j,k}\psi_{j,k}(x),\tag{5.4.4}$$

亦即

$$\lim_{N,M\to+\infty}\left\|f(\circ)-\sum_{j=-M}^{+M}\sum_{k=-N}^{+N}c_{j,k}\psi_{j,k}(\circ)\right\|_{L^2(\mathbb{R})}=0,\tag{5.4.5}$$

则称 $\{\psi_{j,k}\}_{j,k\in\mathbb{Z}}$ 是 $L^2(\mathbb{R})$ 中的一个标准直交基,称 ψ 为一个**直交基小波(orthogonal base wavelet)**,或**离散基小波(discrete base wavelet)**. 称级数 $\sum\limits_{j,k=-\infty}^{+\infty}c_{j,k}\psi_{j,k}(x)$ 为**离散小波级数(discrete wavelet series)**,称

$$c_{j,k}=\langle f,\psi_{j,k}\rangle=\int_{\mathbb{R}}f(x)\overline{\psi_{j,k}(x)}\mathrm{d}x,\quad j,k\in\mathbb{Z}\tag{5.4.6}$$

为离散小波级数的**系数(coefficients of discrete wavelet series)**.

2. 稳定条件、最稳定条件

定义 5.4.5(稳定条件、最稳定条件) 若函数 $\psi\in L^2(\mathbb{R})$ 满足：存在常数 $A,B,0<A\leqslant B<+\infty$，使得 $\forall\omega\in\mathbb{R}$，有

$$A\leqslant\sum_{j=-\infty}^{+\infty}|\hat{\psi}(2^j\omega)|^2\leqslant B,\quad \text{a.e.}\,\omega\in\mathbb{R},$$

则称上述条件为**稳定条件(stability condition)**；若 $0<A=B<+\infty$，则称为**最稳定条件**.

我们有稳定条件的等价性定理.

定理 5.4.7 ψ 的稳定条件等价于：存在常数 $A,B,0<A\leqslant B<+\infty$，使得

$$A\|f\|_{L^2(\mathbb{R})}^2\leqslant\sum_{j=-\infty}^{+\infty}\|(W_\psi f)(j,\circ)\|_{L^2(\mathbb{R})}^2\leqslant B\|f\|_{L^2(\mathbb{R})}^2$$

对每个 $f\in L^2(\mathbb{R})$ 成立，其中 $W_\psi(f)$ 由(5.4.3)式给出.

定理 5.4.8(离散基小波的性质) 设 ψ 是一个满足稳定条件的离散基小波，则它是一个满足容许条件的基小波，且存在常数 $A,B,0<A\leqslant B<+\infty$，使得

$$A\ln 2\leqslant\int_{(0,+\infty)}\frac{|\hat{\psi}(\omega)|^2}{\omega}\mathrm{d}\omega,\quad \int_{(0,+\infty)}\frac{|\hat{\psi}(-\omega)|^2}{\omega}\mathrm{d}\omega\leqslant B\ln 2.$$

证 对积分 $\int_{(1,2)}\frac{|\hat{\psi}(2^j\omega)|^2}{\omega}\mathrm{d}\omega$ 作变量代换，得到

$$\int_{(1,2)}\frac{|\hat{\psi}(2^j\omega)|^2}{\omega}\mathrm{d}\omega=\int_{(2^{-j},2^{-j+1})}\frac{|\hat{\psi}(t)|^2}{t}\mathrm{d}t.$$

另一方面，在稳定条件上除以 ω，得

$$\frac{A}{\omega}\leqslant\sum_{j=-\infty}^{+\infty}\frac{|\hat{\psi}(2^{-j}\omega)|^2}{\omega}\leqslant\frac{B}{\omega},$$

将上式在 $[1,2]$ 上对 ω 积分，得到

$$\int_{[1,2]}\frac{A}{\omega}\mathrm{d}\omega\leqslant\sum_{j=-\infty}^{+\infty}\int_{[1,2]}\frac{|\hat{\psi}(2^{-j}\omega)|^2}{\omega}\mathrm{d}\omega\leqslant\int_{[1,2]}\frac{B}{\omega}\mathrm{d}\omega,$$

即 $A\ln 2\leqslant\sum_{j=-\infty}^{+\infty}\int_{[2^{-j},2^{-j+1}]}\frac{|\hat{\psi}(\omega)|^2}{\omega}\mathrm{d}\omega\leqslant B\ln 2$，于是

$$A\ln 2\leqslant\int_{(0,+\infty)}\frac{|\hat{\psi}(\omega)|^2}{\omega}\mathrm{d}\omega\leqslant B\ln 2.$$

另一个不等式可类似证明.

3. 离散小波变换

定义 5.4.6(离散小波变换) 设离散基小波 ψ 满足稳定条件，对于函数 $f\in L^2(\mathbb{R})$，称

$$(W_\psi f)\left(\frac{1}{2^j},\frac{k}{2^j}\right)=\int_{\mathbb{R}}f(x)\,\overline{\psi_{j,k}(x)}\mathrm{d}x=\langle f,\psi_{j,k}\rangle,\quad j,k\in\mathbb{Z}\tag{5.4.7}$$

为**离散小波变换(discrete wavelet transform)**,或**二进小波变换(dyadic wavelet transform)**.

与(5.4.6)式相比较,离散小波变换就是函数 $f\in L^2(\mathbb{R})$ 按照 $\{\psi_{j,k}\}_{j,k}\in\mathbb{Z}$ 展开成离散小波级数的系数. 在(5.4.3)式 $(W_\psi f)(b,a)=\frac{1}{\sqrt{|a|}}\int_{\mathbb{R}} f(t)\overline{\psi\left(\frac{t-b}{a}\right)}\mathrm{d}t\equiv\langle f,\psi^{ba}\rangle$ 中,令 $b=\frac{1}{2^j}, a=\frac{k}{2^j}, j,k\in\mathbb{Z}$,做适当的变换后,就可得到定义 5.4.6 中的形式.

4. 离散小波变换的反演公式

小波变换的反演问题与 Fourier 变换的反演同样重要. 离散小波变换反演公式如下:

定理 5.4.9 设离散基小波 ψ 满足稳定条件,对于函数 $f\in L^2(\mathbb{R})$,其离散小波变换

$$(W_\psi f)\left(\frac{1}{2^j},\frac{k}{2^j}\right)=\int_{\mathbb{R}} f(x)\,\overline{\psi_{j,k}(x)}\mathrm{d}x=\langle f,\psi_{j,k}\rangle,\quad j,k\in\mathbb{Z}$$

满足反演公式

$$f(x)=\sum_{j,k=-\infty}^{+\infty} d_{j,k}\,\tilde{\psi}_{j,k}(x),\quad x\in\mathbb{R},\tag{5.4.8}$$

其中 $d_{j,k}=\langle f,\tilde{\psi}_{j,k}\rangle=\int_{\mathbb{R}} f(x)\,\overline{\tilde{\psi}_{j,k}(x)}\mathrm{d}x, j,k\in\mathbb{Z}$,且 $\tilde{\psi}_{j,k}(x)=2^{\frac{j}{2}}\tilde{\psi}(2^j x-k), j,k\in\mathbb{Z}$,这里 $\tilde{\psi}$ 是 ψ 的"对偶",由它们的 Fourier 变换确定,即 $(\tilde{\psi})^?(\omega)=\frac{\psi^?(\omega)}{\sum_{k=-\infty}^{+\infty}|\psi^?(\omega+2k\pi)|^2}$.

5.4.4 小波变换应用概述

小波变换有广泛的应用,例如在多分辨分析、编码数据压缩、信噪分离与滤波等方面. 这里不做详细介绍,只给出以下简单叙述.

1. 多分辨分析

定义 5.4.7(多分辨分析) 一个闭子空间序列 $\{V_j\}_{j\in\mathbb{Z}}, V_j\subset L^2(\mathbb{R}), j\in\mathbb{Z}$,满足如下条件:

(1) $V_j\subset V_{j-1}, \forall j\in\mathbb{Z}$;

(2) $\overline{\bigcup_{j\in\mathbb{Z}} V_j}=L^2(\mathbb{R}), \overline{\bigcap_{j\in\mathbb{Z}} V_j}=\{0\}$;

(3) $u(x)\in V_j \Leftrightarrow u(2x)\in V_{j-1}, j\in\mathbb{Z}$;

(4) $u(x)\in V_0 \Rightarrow u(x-k)\in V_0, k\in\mathbb{Z}$;

(5) $\exists g(x)\in V_0$, s. t. $\{g(x-k):k\in\mathbb{Z}\}$ 是 V_0 的 Riesz 基,亦即

$$\forall u(x)\in V_0, \exists!\{\alpha_k:k\in\mathbb{Z}\}\subset l^2,\ \text{s. t. } u(x)=\sum_{k\in\mathbb{Z}}\alpha_k g(x-k),$$

反之，$\forall \{\alpha_k : k \in \mathbb{Z}\} \subset l^2$，$\exists A, B, 0 < A \leqslant B < +\infty$，使得

$$A \| u \|_{L^2(\mathbb{R})}^2 \leqslant \sum_{k \in \mathbb{Z}} |\alpha_k|^2 \leqslant \| u \|_{L^2(\mathbb{R})}^2, \quad \forall u(x) \in V_0,$$

则称$\{V_j\}_{j\in\mathbb{Z}}$为**多分辨分析序列**(**multi-resolution analysis sequence**)，并称 $g(x)$为**生成子**(**generator**).

从 $g(x) \in V_0$ 出发，构造一个**尺度函数**(**scale function**)$\varphi(x) \in V_0$，使得$\{\varphi(x-k)\}_{k\in\mathbb{Z}}$是 V_0 的直交基，例如，可令 $\varphi(x) = \frac{1}{\sqrt{2\pi}}\left(\left(\sum_{k\in\mathbb{Z}} |g^{?}(\omega + 2k\pi)|^2\right)^{-1/2} g^{?}(\omega)\right)^{?}$，则有

$$\varphi(x) \in V_0 \longrightarrow \{\varphi(x-k) : k \in \mathbb{Z}\} \longrightarrow \{\varphi_{jk}(x) = 2^{-j/2}\varphi(2^{-j}x - k) : j, k \in \mathbb{Z}\}$$

↑ 尺度函数　　↑ V_0 的直交基　　↑ V_j 的直交基

具体地，对于给定的多分辨分析序列$\{V_j\}_{j\in\mathbb{Z}}$，取生成子 $g(x) \in V_0$，生成一个尺度函数 $\varphi(x) \in V_0$，使得 $\{\varphi(x-k) : k \in \mathbb{Z}\}$ 是 V_0 的直交基. 于是$\frac{1}{\sqrt{2}}\varphi\left(\frac{x}{2}\right) \in V_1 \subset V_0$.

令

$$\frac{1}{\sqrt{2}}\varphi\left(\frac{x}{2}\right) = \sum_{k\in\mathbb{Z}} h_k \varphi(x-k), \tag{5.4.9}$$

$$\varphi_{jk}(x) = 2^{-\frac{j}{2}}\varphi(2^{-j}x - k), \quad j, k \in \mathbb{Z}, \tag{5.4.10}$$

其中

$$h_k = \left\langle \frac{1}{\sqrt{2}}\varphi\left(\frac{x}{2}\right), \varphi(x-k) \right\rangle = \frac{1}{\sqrt{2}}\int_{\mathbb{R}} \varphi\left(\frac{x}{2}\right)\overline{\varphi}(x-k)\mathrm{d}x. \tag{5.4.11}$$

对 φ 取 Fourier 变换，得到

$$\varphi^{?}(2\omega) = H(\omega)\varphi^{?}(\omega), \tag{5.4.12}$$

其中

$$H(\omega) = \frac{1}{\sqrt{2}}\sum_{k\in\mathbb{Z}} h_k \mathrm{e}^{-\mathrm{i}k\omega}, \tag{5.4.13}$$

且 $H(\omega) \in L^2([0, 2\pi])$.

我们称$\{h_k\}_{k\in\mathbb{Z}}$为**频率响应**(**response of frequency**)，称 $H(\omega)$为$\{h_k\}_{k\in\mathbb{Z}}$的**传递函数**(**transfer function**).

定理 5.4.10 对于多分辨分析序列$\{V_j\}_{j\in\mathbb{Z}}$的生成子 $\varphi(x)$，频率响应$\{h_k\}_{k\in\mathbb{Z}}$与传递函数 $H(\omega)$，满足

(1) $|H(\omega)|^2 + |H(\omega+\pi)|^2 = 1$；

(2) 若$\{h_k\}_{k\in\mathbb{Z}} \in l^1$，且 $\varphi^{?}(\omega)$连续，$\varphi^{?}(0) = 1$，则 $H(0) = 1$.

2. 多分辨分析的直交补

在确定多分辨分析序列$\{V_j\}_{j\in\mathbb{Z}}$之后,我们定义一组新的闭子空间 $W_j, j\in\mathbb{Z}$,使得(图 5.4.1)

$$V_j \oplus W_j = V_{j-1}, \quad j\in\mathbb{Z},$$

于是,W_j 是 V_j 的直交补,且有 $V_j\oplus W_j\oplus W_{j-1}\oplus\cdots\oplus W_{j-m+1}=V_{j-m}$,以及

$$\bigoplus_{j=-\infty}^{J} W_j = V_{J-1},\ J\in\mathbb{Z};\quad V_J\oplus\bigoplus_{j=J}^{+\infty} W_j = L^2(\mathbb{R}^n),\ J\in\mathbb{Z};\quad \bigoplus_{j=-\infty}^{+\infty} W_j = L^2(\mathbb{R}^n).$$

直交补$\{W_j\}_{j\in\mathbb{Z}}$具有如下的性质.

定理 5.4.11 对于直交补$\{W_j\}_{j\in\mathbb{Z}}$,有

(1) $u(x)\in W_j \Rightarrow u(x-2^jk)\in W_j,\ j,k\in\mathbb{Z}$;

(2) $u(x)\in W_j \Leftrightarrow u(2x)\in W_{j-1},\ j\in\mathbb{Z}$;

(3) $\mathrm{P}_{W_j}u(x)\to 0, |j|\to+\infty, \forall u\in L^2(\mathbb{R}^n)$,

这里 $\mathrm{P}_{W_j}u$ 是 $u\in L^2(\mathbb{R}^n)$在空间 W_j 上的直交投影.

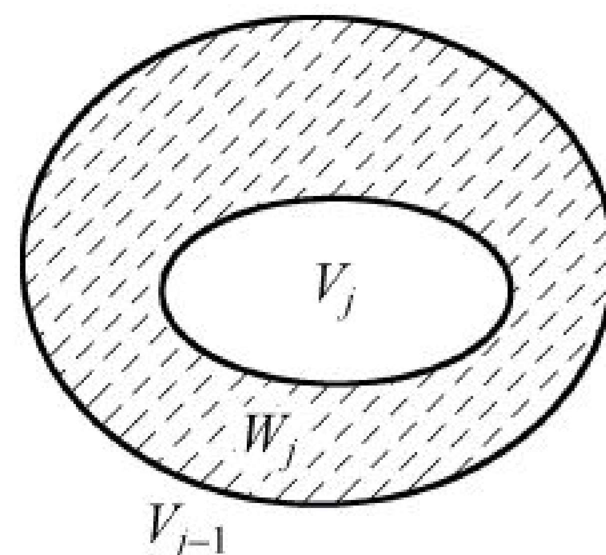


图 5.4.1 直交补

做多分辨分析序列$\{V_j\}_{j\in\mathbb{Z}}$的直交补$\{W_j\}_{j\in\mathbb{Z}}$,在直交补中,同样寻找其标准直交系,进行信号分析时,将信号在$\{V_j\}_{j\in\mathbb{Z}}$与$\{W_j\}_{j\in\mathbb{Z}}$中进行分解.

3. 小波函数 ψ

我们希望从尺度函数 φ 出发,构造一个小波函数 $\psi\in W_0$,使得对于每个 $j\in\mathbb{Z}$,集$\{\psi_{j,k}:k\in\mathbb{Z}\}$是 W_j 的直交基,其中 $\psi_{j,k}(x)=2^{-j/2}\psi(2^{-j}x-k)$.

由 φ 出发,令

$$\psi(x) = \sum_{k\in\mathbb{Z}} g_k\,\varphi_{-1,k}(x) = 2^{1/2}\sum_{k\in\mathbb{Z}} g_k\,\varphi(2x-k), \tag{5.4.14}$$

其中 $\varphi_{-1,k}$取自 $\varphi_{jk}(x)=2^{-j/2}\varphi(2^{-j}x-k), j,k\in\mathbb{Z}$,即

$$\varphi_{-1,k}(x) = 2^{1/2}\varphi(2x-k),\quad k\in\mathbb{Z},$$

且 $g_k=\langle\psi,\varphi_{-1,k}\rangle=\langle\psi,2^{1/2}\varphi(2x-k)\rangle$. 如此得到的由(5.4.14)式决定的 ψ,称为**小波函数**.

小波函数 ψ 具有如下性质:

(1) $\frac{1}{\sqrt{2}}\psi\left(\frac{x}{2}\right)=\sum_{k\in\mathbb{Z}} g_k\varphi(x-k)$,其中 $g_k=\left\langle\frac{1}{\sqrt{2}}\psi\left(\frac{x}{2}\right),\varphi(x-k)\right\rangle$;

$\left(\text{与(5.4.7)式}\frac{1}{\sqrt{2}}\varphi\left(\frac{x}{2}\right)=\sum_{k\in\mathbb{Z}} h_k\varphi(x-k)\ \text{相比较}, h_k=\left\langle\frac{1}{\sqrt{2}}\varphi\left(\frac{x}{2}\right),\varphi(x-k)\right\rangle.\right)$

(2) $\hat{\psi}(2\omega)=G(\omega)\hat{\varphi}(\omega)$，其中 $G(\omega)=\frac{1}{\sqrt{2}}\sum\limits_{k\in\mathbf{Z}}g_k e^{-ik\omega}\in L^2([0,2\pi])$.

$\left(\text{与(5.4.12)式 } \hat{\varphi}(2\omega)=H(\omega)\hat{\varphi}(\omega)\text{ 相比较，} H(\omega)=\frac{1}{\sqrt{2}}\sum\limits_{k\in\mathbf{Z}}h_k e^{-ik\omega}.\right)$

下面的(图 5.4.2)是上述建立过程的图解，可以直接将其付诸应用.

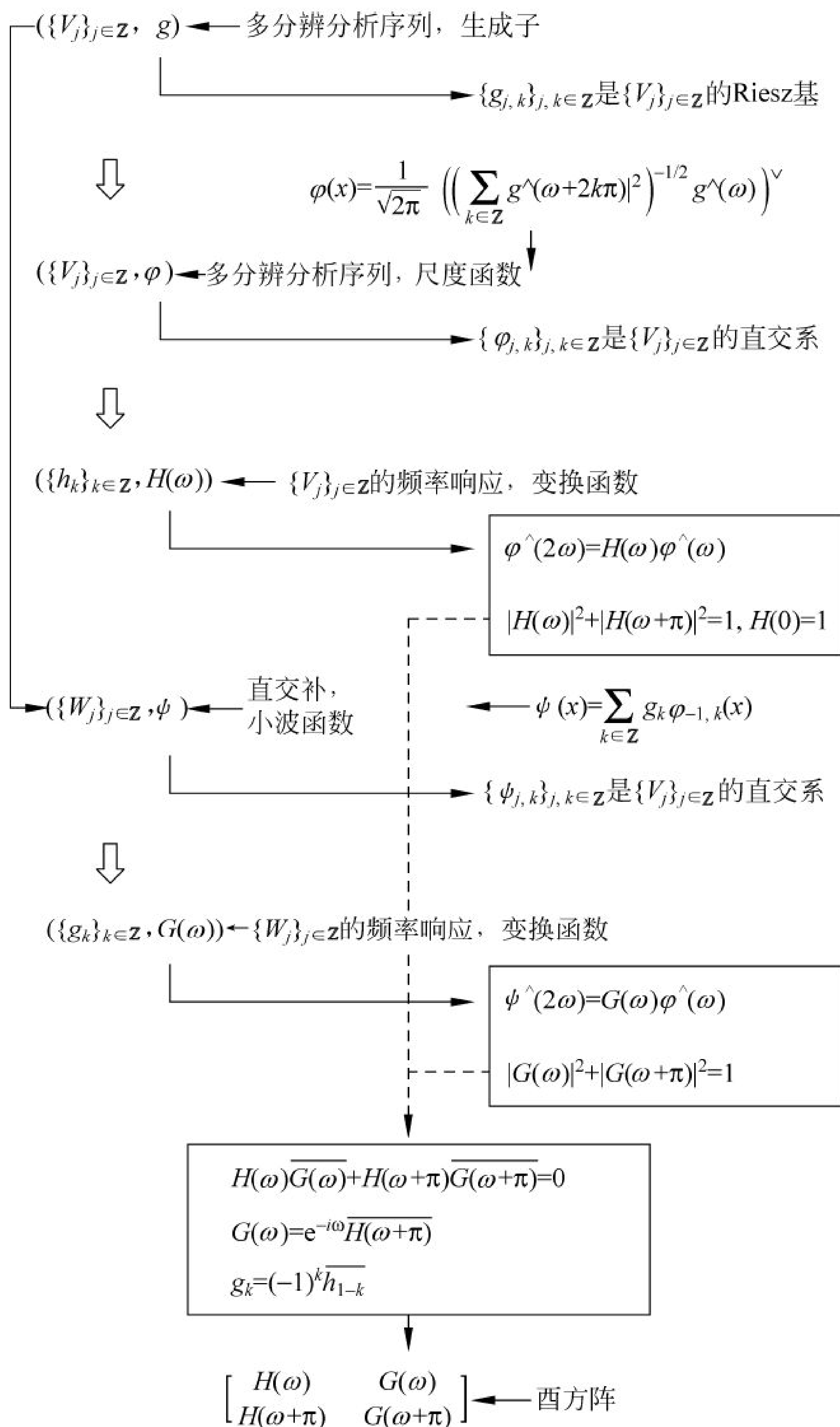


图 5.4.2 小波应用概述

由尺度函数 $\varphi(x)$ 得到的小波函数 $\psi(x)$ 与相应的 $G(\omega)$，满足

(1) $\psi\in W_0 \Leftrightarrow H(\omega)\overline{G(\omega)}+H(\omega+\pi)\overline{G(\omega+\pi)}=0$；

(2) $\{\psi(x-k)\}_{k\in\mathbb{Z}}$是直交系 $\Leftrightarrow |G(\omega)|^2+|G(\omega+\pi)|^2=1$；

(3) $\{\psi(x-k)\}_{k\in\mathbb{Z}}$是 W_0 的直交系 $\Leftrightarrow |G(\omega)|^2+|G(\omega+\pi)|^2=1$ 与

$$H(\omega)\overline{G(\omega)}+H(\omega+\pi)\overline{G(\omega+\pi)}=0$$

成立．于是

$$(\{V_j\}_{j\in\mathbb{Z}},g)\Rightarrow(\{V_j\}_{j\in\mathbb{Z}},\varphi)\Rightarrow(\{W_j\}_{j\in\mathbb{Z}},\psi).$$

4. 信号分析，Mallat 算法

对于随时间 t 改变的信号 $f(t)$，下面做信号分析．采用上面图 5.4.2 中的记号．

设 $f\in V_{J_1}$，$J_1\in\mathbb{Z}$ 为固定的整数．令

$$f(t)\equiv A_{J_1}f(t)=\sum_{k\in\mathbb{Z}}c_{J_1,k}\,\varphi_{J_1,k}(t), \tag{5.4.15}$$

其中 $\langle\varphi_{J_1,k},\varphi_{J_1+1,m}\rangle=\overline{h_{k-2m}}$，$\langle\varphi_{J_1,k},\psi_{J_1+1,m}\rangle=\overline{g_{k-2m}}$，$k,m\in\mathbb{Z}$．

将 $f(t)$ 分解为

$$f(t)=A_{J_1}f(t)=A_{J_1+1}f(t)+B_{J_1+1}f(t), \tag{5.4.16}$$

其中 $A_{J_1+1}f(t)=\sum\limits_{m\in\mathbb{Z}}C_{J_1+1,m}\,\varphi_{J_1+1,m}(t)$ 具有系数

$$C_{J_1+1,m}=\sum_{k\in\mathbb{Z}}\overline{h_{k-2m}}C_{J_1,k}; \tag{5.4.17}$$

而 $B_{J_1+1}f(t)=\sum\limits_{m\in\mathbb{Z}}D_{J_1+1,m}\,\psi_{J_1+1,m}(t)$ 有系数

$$D_{J_1+1,m}=\sum_{k\in\mathbb{Z}}\overline{g_{k-2m}}C_{J_1,k}. \tag{5.4.18}$$

于是，信号在多分辨分析序列 $\{V_j\}_{j\in\mathbb{Z}}$ 与其直交补 $\{W_j\}_{j\in\mathbb{Z}}$ 中得到分解．再将(5.4.17)式、(5.4.18)式用新记号表示为

$$C_{J_1+1}\equiv C_{J_1+1,m},\quad \widetilde{H}C_{J_1}\equiv\sum_{k\in\mathbb{Z}}\overline{h_{k-2m}}C_{J_1,k},$$

$$B_{J_1+1}\equiv D_{J_1+1,m},\quad \widetilde{G}C_{J_1}\equiv\sum_{k\in\mathbb{Z}}\overline{g_{k-2m}}C_{J_1,k},$$

得到 $C_{J_1+1}=\widetilde{H}C_{J_1}$，$B_{J_1+1}=\widetilde{G}C_{J_1}$，其中 $\widetilde{H}=[H_{m,k}]=[\overline{h}_{k-2m}]$，$\widetilde{G}=[G_{m,k}]=[\overline{g}_{k-2m}]$是两个无穷矩阵．于是

$$f(t)=A_{J_2}f(t)+\sum_{j=J_1+1}^{J_2}B_jf(t), \tag{5.4.19}$$

$$A_jf(t)=\sum_{k\in\mathbb{Z}}C_{j,k}\,\varphi_{j,k}(t),\quad B_jf(t)=\sum_{k\in\mathbb{Z}}D_{j,k}\,\psi_{j,k}(t), \tag{5.4.20}$$

$$C_{j+1}=\widetilde{H}C_j,\ B_{j+1}=\widetilde{G}C_j,\quad j=J_1,J_1+1,\cdots,J_2-1, \tag{5.4.21}$$

(5.4.20)式与(5.4.21)式称为信号 $f(t)$ 的 **Mallat 分解公式**，称

(1) $A_j f$ 为 $f(t)$在 2^j 分辨率之下的连续近似(continuous approximation),是 $f(t)$的频率在 2^{-j}之下的部分,C_j 为 $f(t)$在 2^j 分辨率之下的离散近似(discrete approximation);

(2) $B_j f$ 为 $f(t)$在 2^j 分辨率之下的连续细节(continuous details),是 $f(t)$的频率在 2^{-j}与 2^{-j+1}之间部分,D_j 为 $f(t)$在 2^j 分辨率之下的离散细节(discrete details).

若设$\widetilde{H}^*$,$\widetilde{G}^*$分别是$\widetilde{H}$,$\widetilde{G}$在空间 l^2 的共轭算子,或说$\widetilde{H}^*$,$\widetilde{G}^*$分别是$\widetilde{H}$,$\widetilde{G}$的共轭变换.于是,一个信号 $f(t)$可以被分解为不同的频率分量 $A_j f$,$D_j f$,根据问题的要求而确定高频部分、细节部分.则可将小波的应用归纳为如下框图(图 5.4.3):

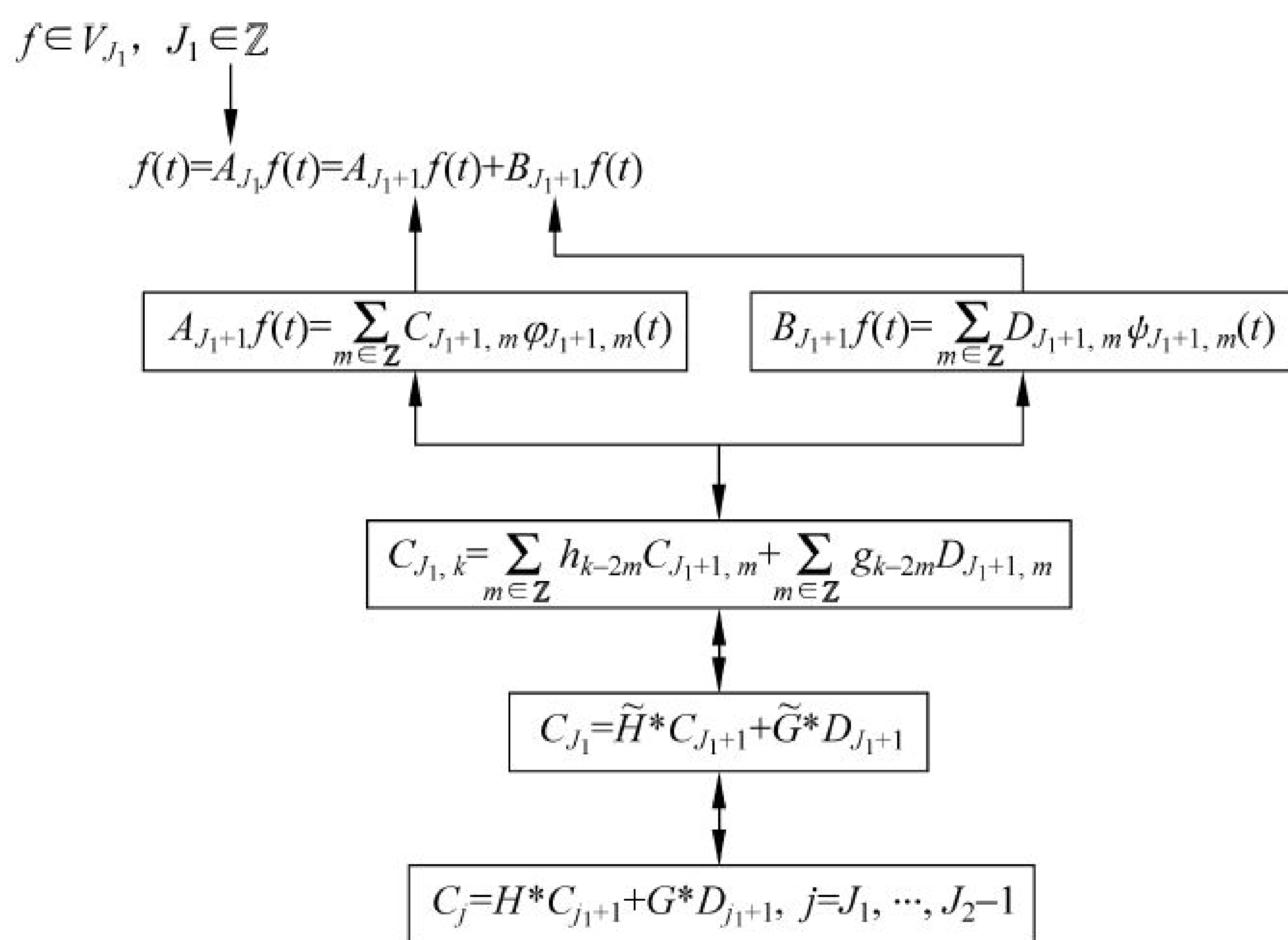


图 5.4.3 小波应用框图(1)

5. 编码数据压缩

输入信号 $f(t)$,根据物理仪器的分辨率(为有限数),令

$$f(t)\in V_{J_1},\tag{5.4.22}$$

$$f(t)=A_{J_1}f(t)+\sum_{k\in\mathbb{Z}}C_{J_1,k}\varphi_{J_1,k}(t),\tag{5.4.23}$$

其中$\{V_j\}_{j\in\mathbb{Z}}$是 $L^2(\mathbb{R})$中的多分辨分析序列的闭子空间序列,$V_j\subset L^2(\mathbb{R})$,$j\in\mathbb{Z}$,系数$\{C_{J_1,k}\}_{k\in\mathbb{Z}}$具有有限长度(即有限项).进而,为复原信号,只需用递归算法

$$C_j=H^*C_{j+1}+G^*D_{j+1},\quad j=J_2-1,\cdots,J_1.$$

根据 Mallat 算法,将 $f(t)$分解为不同的频率(即放入不同的频道).

应用时,对$\{h_k\}$,$\{g_k\}$的长度提出要求,也要求长度$\ll N_{J_1}$,于是,当 C_{J_1},D_{J_1} 减半时,C_j,D_j 均为$\dfrac{N_{J_1}}{2^{j-J_1}}$,$j=J_1+1,\cdots,J_2$,从而达到数据压缩的效果.我们归纳为如下框图(图 5.4.4).

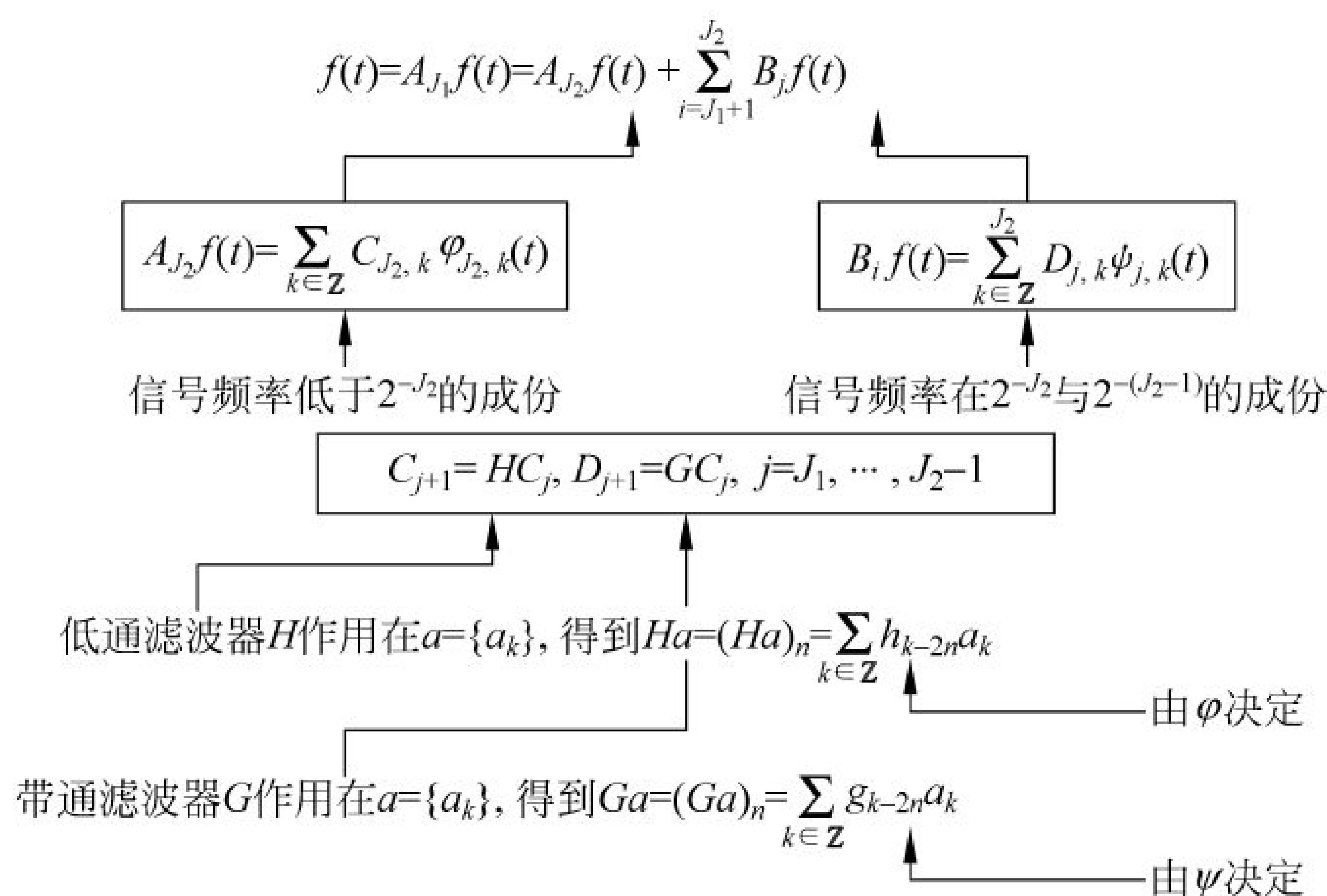


图 5.4.4 小波应用框图(2)

6. 信噪分离与滤波

信噪分离与滤波原理 将信号 $f(t)$分解为低频(频率小于 2^{-J_2})成分与频率介于 2^{-j} 与 $2^{-(j-1)}$ 的成分,$J_1-1<j<J_2$,每个频率成分 $A_{J_2}f(t)$与 $\mathfrak{B}f(t)$又分解成具有相同频率“程区”、但不同位相的成分

$$A_{J_2}f(t)=\sum_{k\in\mathbf{Z}}C_{J_2,k}\varphi_{J_2,k}(t) \quad 与 \quad \mathfrak{B}f(t)=\sum_{k\in\mathbf{Z}}^{J_2}D_{j,k}\psi_{j,k}(t).$$

分析 如果尺度函数与小波函数的能量分别集中在 $t=a$ 与 $t=b$ 附近,则

$$A_{J_2}f(t)=\sum_{k\in\mathbf{Z}}C_{J_2,k}\varphi_{J_2,k}(t),\quad B_jf(t)=\sum_{k\in\mathbf{Z}}^{J_2}D_{j,k}\psi_{j,k}(t).$$

$C_{J_2,k}\varphi_{J_2,k}(t)$与 $D_{j,k}\psi_{j,k}(t)$的能量分别集中在 $t=2^j(k+a)$与 $t=2^j(t+b)$附近. 由此可见,时间(空间)域中的分辨率随频率的大小而调节,高频处精密、低频处粗疏.

由 Mallat 分解算法,得到

$$C_{j+1}=HC_j,\quad D_{j+1}=GC_j,\quad j=J_1,\cdots,J_2-1.$$

将信号 $f(t)$按照图 5.4.4 分解,再根据一定的先验知识,区分有效信号与噪声,将与噪声相应的 $C_{J_2,k},D_{j,k},j=J_1+1,\cdots,J_2$置零,得到一个调整后的新序列 $\widetilde{C}_{J_2,k},\widetilde{D}_{j,k},j=J_1+1,\cdots,J_2$;并且按照

$$\widetilde{C}_{j-1}=H^*\widetilde{C}_j+G^*\widetilde{D}_j,\quad j=J_1+1,\cdots,J_2,$$

最后得到去掉噪声后的信号

$$\tilde{f}(t)=A_{J_1}\tilde{f}(t)=\sum_{k\in\mathbf{Z}}\widetilde{C}_{J_1,k}\varphi_{J_1,k}(t).$$

有关具体问题的小波方法,可以通过实际问题去熟悉.

习题 5

1. 试证：$\mathbb{S}(\mathbb{R}^n)\subset L^p(\mathbb{R}^n)$，$1\leqslant p<+\infty$.

2. 试证：$e^{-x^2}\in\mathbb{S}(\mathbb{R}^n)$.

3. 试证：$|x^\beta\partial^p\varphi(x)|\leqslant M_{\beta,p},\forall\beta,p\in\mathbb{N}^n\Leftrightarrow|(1+|x|^2)^{\beta/2}\partial^p\varphi(x)|\leqslant M_{\beta,p},\forall\beta,p\in\mathbb{N}^n\Leftrightarrow\lim\limits_{|x|\to+\infty}|x^\beta\partial^p\varphi(x)|=0,\forall\beta,p\in\mathbb{N}^n$.

4. 将 Schwartz 分布的各种运算推广到 $E^*(\mathbb{R}^n)$ 与 $D^*(\mathbb{R}^n)$.

5. 试证：在分布意义下，$\mathbb{S}^*(\mathbb{R}^n)$，$E^*(\mathbb{R}^n)$，$D^*(\mathbb{R}^n)$ 中成立等式 $\dfrac{\partial^2u}{\partial x_i\partial x_j}=\dfrac{\partial^2u}{\partial x_j\partial x_i}$.

6. $\forall a\in C^\infty(\mathbb{R}^n)$，$u\in D^*(\mathbb{R}^n)$，证明等式 $\dfrac{\partial}{\partial x_i}(au)=a\dfrac{\partial u}{\partial x_i}+\dfrac{\partial a}{\partial x_i}u$ 在分布意义下成立.

7. 试证包含关系 $D^*(\mathbb{R}^n)\supset\mathbb{S}^*(\mathbb{R}^n)\supset E^*(\mathbb{R}^n)$.

8. 试证 $\operatorname{supp}\delta=\{0\}$，$\delta\in E^*(\mathbb{R}^n)$.

9. 设 $f\in L^1(\mathbb{R})$，试证 $f^?(\xi)$ 是 $\mathbb{R}$ 上的偶函数，当且仅当 $f(x)$ 是 $\mathbb{R}$ 上的偶函数，并证明 $[f(-\circ)]^?(\xi)=f^?(-\xi)$.

10. 试证定理 5.2.1、定理 5.2.7 以及定理 5.2.8.

11. 利用定理 5.2.4，证明 $(L^1([a,b]),+,\alpha\cdot,\|f\|_{L^1([a,b])},f*g)$ 是一个没有单位元的 Banach 代数.

12. 对于 $L^2(\mathbb{R})$ 的 Fourier 变换，证明相应于 $L^1(\mathbb{R})$ 上 Fourier 变换的性质.

13. 对于 $L^p(\mathbb{R}^n)(1<p<2)$ 的 Fourier 变换，证明相应于 $L^p(\mathbb{R}^n)$ 的 Fourier 变换的性质.

14. 能否将定义 $L^p(\mathbb{R}^n)(1<p<2)$ 上 Fourier 变换的方法推广到 $p>2$？为什么？

15. 证明 Fourier 变换的惟一性定理.

16. 设 $\Omega_1\subset\mathbb{R}^n$，$\Omega_2\subset\mathbb{R}^m$ 为开集，$K_1\subset\mathbb{R}^n$，$K_2\subset\mathbb{R}^m$ 为紧致集，$\varphi\in D(\Omega_1\times\Omega_2)$，且 $\operatorname{supp}\varphi\in K_1\times K_2$，试证：$\forall T\in D^*(\Omega_1\times\Omega_2)$，成立 $\int\langle u,\varphi(\circ,y)\rangle\mathrm{d}y=\left\langle u,\int\varphi(\circ,y)\mathrm{d}y\right\rangle$.

17. 设 $\Omega_1\subset\mathbb{R}^n$，$\Omega_2\subset\mathbb{R}^m$ 为开集，若 $f\in D(\Omega_1\times\Omega_2)$，试证积分和 $\sum\limits_j f(\xi_j,y)\Delta_jx$ 在 $D(\Omega_2)$ 中收敛于 $\int f(x,y)\mathrm{d}x$.

18. 设 $S\in E^*(\mathbb{R})$，$T\in E(\mathbb{R})$，定义 $S\in E^*(\mathbb{R})$ 的 Laplace 变换 $L_S(\lambda)=\langle S(\cdot),e^{-\lambda\cdot}\rangle$，试证 $S*T$ 的 Laplace 变换 $L_{S*T}(\lambda)=L_S(\lambda)\cdot L_T(\lambda)$.

19. 设 $S\in D^*(\mathbb{R}^n)$，$\varphi\in D(\mathbb{R}^n)$，试证 $\langle S,\varphi\rangle\to S*\varphi$ 是双线性映射，且对 S 与 φ 分别是连续的.

20. 计算卷积：$e^{-|x|}*e^{-|x|}$；$e^{-ax^2}*e^{-ax^2}$，$a>0$；$xe^{-ax^2}*xe^{-ax^2}$，$a>0$.

21. 试证：平移算子 τ_h 与卷积算子 $T*$，$T\in E^*(\mathbb{R}^n)$ 是可交换的，即

$$\tau_h(T*S)=T*\tau_hS,\quad\forall S\in D^*(\mathbb{R}^n);$$

并且证明：$\tau_hT=\delta(x-h)*T$.

22. 试证卷积映射 $(T,S)\to T*S$ 是双线性的，并且对于 T,S 是连续的.

23. 试证：将 $\mathbb{S}^*(\mathbb{R}^n)$ 嵌入 $D^*(\mathbb{R}^n)$，$\mathbb{S}^*(\mathbb{R}^n)\to D^*(\mathbb{R}^n)$，若 $f\in\mathbb{S}^*(\mathbb{R}^n)$ 从零元嵌为零元，则嵌入映射是

单射.

24. 试证：$L^p(\mathbb{R}^n)\subset\mathbb{S}^*(\mathbb{R}^n)$，$1\leqslant p<+\infty$.

25. 在定理 5.3.9(4)中用到了两个分布的 Fourier 变换的乘积 $T^? S^?$. 请根据定理后的注的说明，给出：对于 $T\in\mathbb{S}^*(\mathbb{R}^n)$，$S\in E^*(\mathbb{R}^n)$，乘积 TS 的定义；试证：在定理 5.3.9(4)的假设下，对于 $T\in S^*(\mathbb{R}^n)$，$S\in E^*(\mathbb{R}^n)$，“$\langle T^? S^?,\varphi\rangle=\langle T^?,S^?\varphi\rangle$，$\forall\,\varphi\in\mathbb{S}(\mathbb{R}^n)$”有意义.

26. 计算乘积 $x\delta$ 的 Fourier 变换，其中 δ 是狄拉克分布.

27. 试证定理 5.3.11.

28. 试证例 5.4.1、例 5.4.2、例 5.4.3 中的函数是基小波.

29. 对于 Fourier 分析与小波分析作详细的比较(包括连续的与离散的).

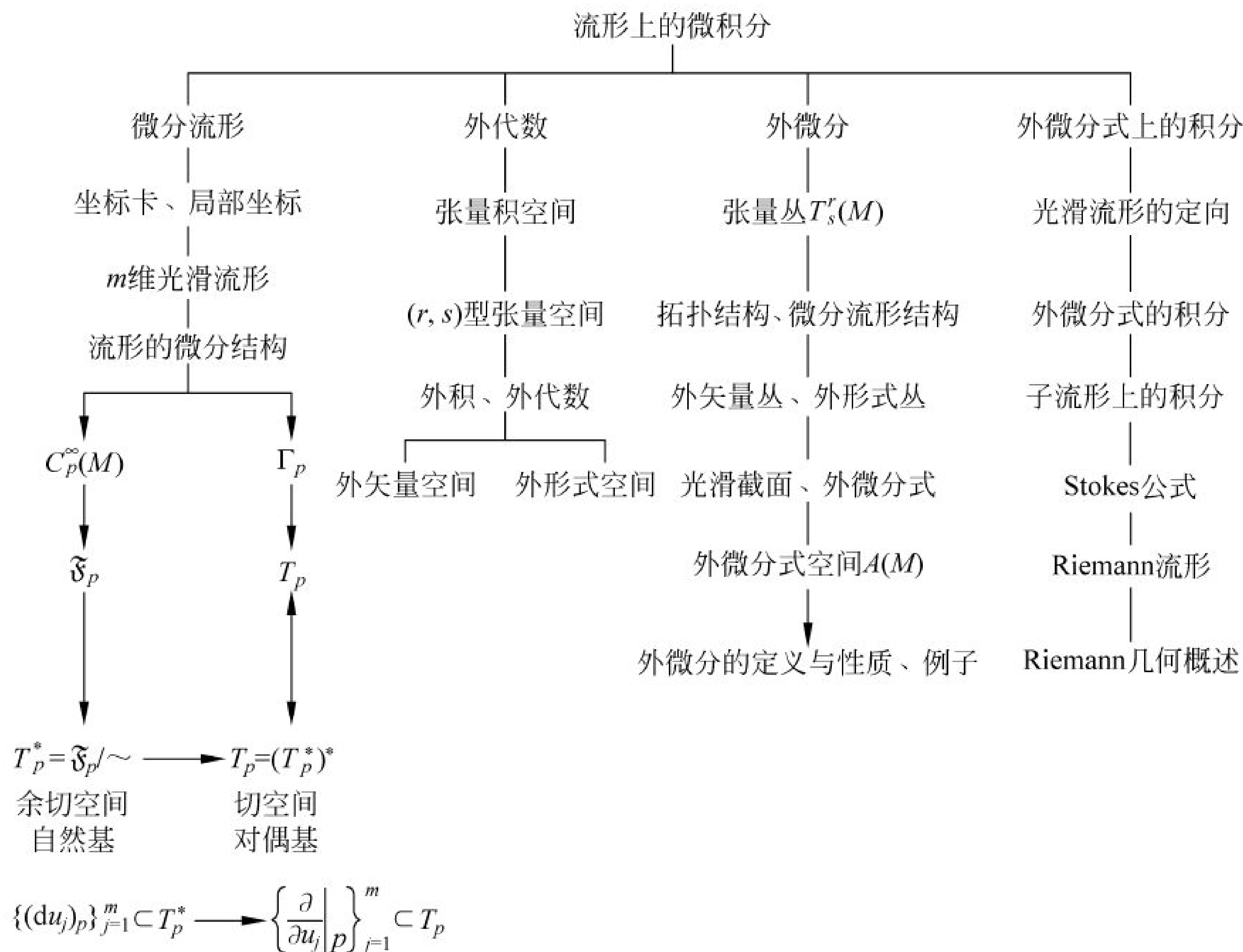
30. 请读者给出一个具体例子，使用小波分析方法来分析一个信号 $f(t)$，并给出它的 Mallat 算法.

第6章 流形上的微积分

数学大师陈省身曾预言：将来的数学研究对象，必然是流形．当今数学主流之一“流形上的分析”已经成为科学领域中诸多学科的必要基础．不仅数学科学，而且物理学、天文学、化学、地质学、气象学、生物学，甚至医学、社会学、人文科学等诸多领域的研究中，都需要关于流形的知识．

流形的研究中蕴含着丰富的现代数学思维方法与技巧，包含了大量的现代科技发展的预见与需求．本章所选内容，部分源自陈省身先生的《微分几何讲义》[3]．

一方面，流形上微积分的起点是经典微积分，经典微积分在几何上的应用演变成“曲线论”与“曲面论”；另一方面，流形上的微积分又作为微分几何的基本知识，置于近代应用数学基础中的重要位置．本章主要参考文献是[3]，[7]，[9]，[15]，[16]．



6.1 基本概念

微分几何的创始人是 C. F. Gauss(1777—1855),他引进了曲面的第一基本形式,建立了曲面理论,奠定了曲面的"弯曲"几何. B. Riemann(1826—1866)则把 Gauss 的理论推广到 n 维空间,从而诞生了 Riemann 几何. Riemann 几何的重要性在于:它成为 Einstein 广义相对论的理论基础,因为 Enistein 把引力现象解释成 Riemann 空间的曲率性质,这就使物理现象有了数学基础,而数学概念也有了现实世界的支撑. 进而,Yang-Mills 的规范场论更是物理与数学结合的典范.

目前,从欧氏几何到非欧几何,从投影几何、仿射几何、微分几何到 Riemann 几何,通过外代数、外微分、流形、曲率、连络、标架、同调,……直到微分拓扑、代数拓扑、几何分析,……形成了严密的、精湛的数学框架,并在自然科学领域中获得了卓有成效的应用.

本章只能介绍流形上的微积分,作为最基础的知识,使得读者达到向更高深理论与应用前进的起点,提高读者的数学素养与自学更深层次知识的能力.

6.1.1 微分流形结构

1. 微分流形

基本集合取为 $\mathbb{R}^m=\{x=(x_1,x_2,\cdots,x_m):x_j\in\mathbb{R},j=1,2,\cdots,m\}$,即 m 维欧氏空间;$\Omega\subseteq\mathbb{R}^m$ 为 $\mathbb{R}^m$ 中的开集;$f:\Omega\to\mathbb{R}$ 为 Ω 上的实值函数. $C(\Omega)\equiv C^0(\Omega)=\{f:\Omega\to\mathbb{R},f$ 是 Ω 上的连续函数$\}$;$C^r(\Omega)=\{f:\Omega\to\mathbb{R},\partial^k f\in C(\Omega),k=0,1,\cdots,r\},r\in\mathbb{N}$;$C^\infty(\Omega)=\{f:\Omega\to\mathbb{R},\partial^k f\in C(\Omega),\forall k=0,1,\cdots\}$.

首先构建 T_2 型拓扑空间(Hausdorff 空间)(X,τ)上的流形结构.

定义 6.1.1(拓扑流形、坐标卡、图册) 设(X,τ)为 Hausdorff 拓扑空间,若 $\forall p\in X$,存在开邻域 $U_p\in\tau$,$p\in U_p$,使得 U_p 与 $\mathbb{R}^m$ 中一个开集同胚,则称 X 为一个 m 维拓扑流形,记 $X=M$,记同胚映射为 $\varphi_{U_p}\equiv\varphi$: $U_p\to\varphi(U_p)$. 易见 $U_p\subset X$,$\varphi(U_p)\subset\mathbb{R}^m$,且映射 φ 是双方单值、双方连续的. 于是,$\varphi_{U_p}(U_p)\subset\mathbb{R}^m$ 是 $\mathbb{R}^m$ 中的开集. 省去下标,记$(U,\varphi_U)\equiv(U_p,\varphi_{U_p})$,称$(U,\varphi_U)$为 M 的**一个坐标卡(chart)**,称 $\mathbb{A}=\{(U,\varphi_U):U\in\tau\}$为**图册(atlas)**.

M 上的坐标卡(U,φ_U)如图 6.1.1 所示.

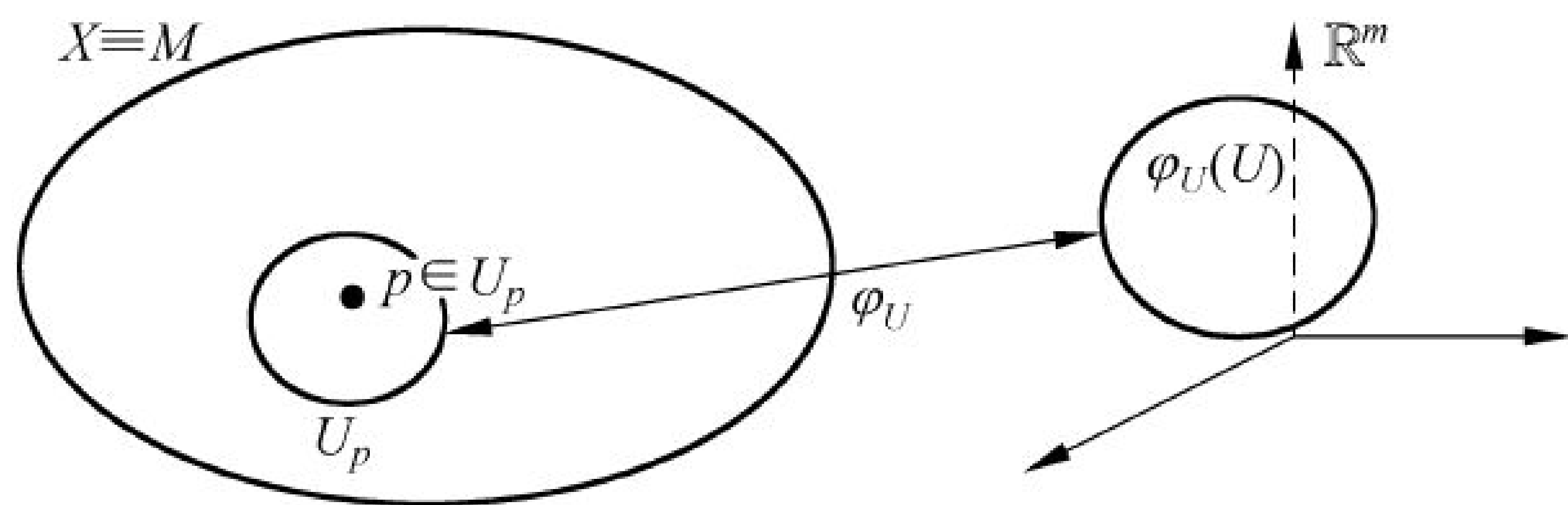


图 6.1.1 M 上的坐标卡

然而,对一个点 $p\in M$,可以有多个 $U\in\tau$,使得对于每个 U,都有 $p\in U$,故 p 点的坐标卡(U,φ_U)不止一个,为此,必须定义坐标卡的"相容性",并得出"相容条件".

设(U,φ_U),(V,φ_V)为流形 M 的两个坐标卡,考虑以下两种情形:

(1) $U\cap V=\varnothing \Rightarrow$ 任一点 $p\in U,p\notin V$,取(U,φ_U)为点 $p\in M$ 的坐标卡;

(2) $U\cap V\neq\varnothing \Rightarrow$ 任一点 $p\in U\cap V$,$\varphi_U(U\cap V)$与 $\varphi_V(U\cap V)$均为$\mathbb{R}^m$ 中的非空开集 $\Rightarrow$ $\varphi_U(p)\in\varphi_U(U\cap V)$,$\varphi_V(p)\in\varphi_V(U\cap V)$ 应当满足复合条件:$\varphi_V\circ\varphi_U^{-1}|_{\varphi_U(U\cap V)}:\varphi_U(U\cap V)\to\varphi_V(U\cap V)$为$\mathbb{R}^m$ 到$\mathbb{R}^m$ 的同胚映射,且其逆映射$(\varphi_V\circ\varphi_U^{-1}|_{\varphi_U(U\cap V)})^{-1}:\varphi_V(U\cap V)\to\varphi_U(U\cap V)$是$\mathbb{R}^m$ 到$\mathbb{R}^m$ 的同胚映射 $\Rightarrow$ 令 $F\equiv\varphi_V\circ\varphi_U^{-1}|_{\varphi_U(U\cap V)}$,它是开集 $\varphi_U(U\cap V)$到开集 $\varphi_V(U\cap V)$的同胚.称满足(1)或(2)的两个坐标卡为相容的.

设 $F(z)=(f_1(z),f_2(z),\cdots,f_m(z))\in\mathbb{R}^m$,$\forall z\in\varphi_U(U\cap V)$,则 $\varphi_U(p)=((\varphi_U(p))_1,(\varphi_U(p))_2,\cdots,(\varphi_U(p))_m)\equiv(u_1,u_2,\cdots,u_m)$. 令 $u_j=(\varphi_U(p))_j,p\in U,j=1,2,\cdots,m$.

定义 6.1.2(局部坐标、M 上的局部坐标系) 对于任一 $p\in U$,$\varphi_U(p)\in\mathbb{R}^m$,称 $u_j=(\varphi_U(p))_j(1\leqslant j\leqslant m)$为 p 的**局部坐标(local coordinates)**,称

$$\{(u_1,u_2,\cdots,u_m):u_j=(\varphi_U(p))_j,\ j=1,2,\cdots,m,\ (U,\varphi_U)\in\mathbb{A}\}$$

为 M 的**局部坐标系(local coordinate system)**,简记为

$$\{(u_j=(\varphi_U(p))_j)_{j=1}^m:(U,\varphi_U)\in\mathbb{A}\},$$

或记为$\{(U,u_j)\}$.

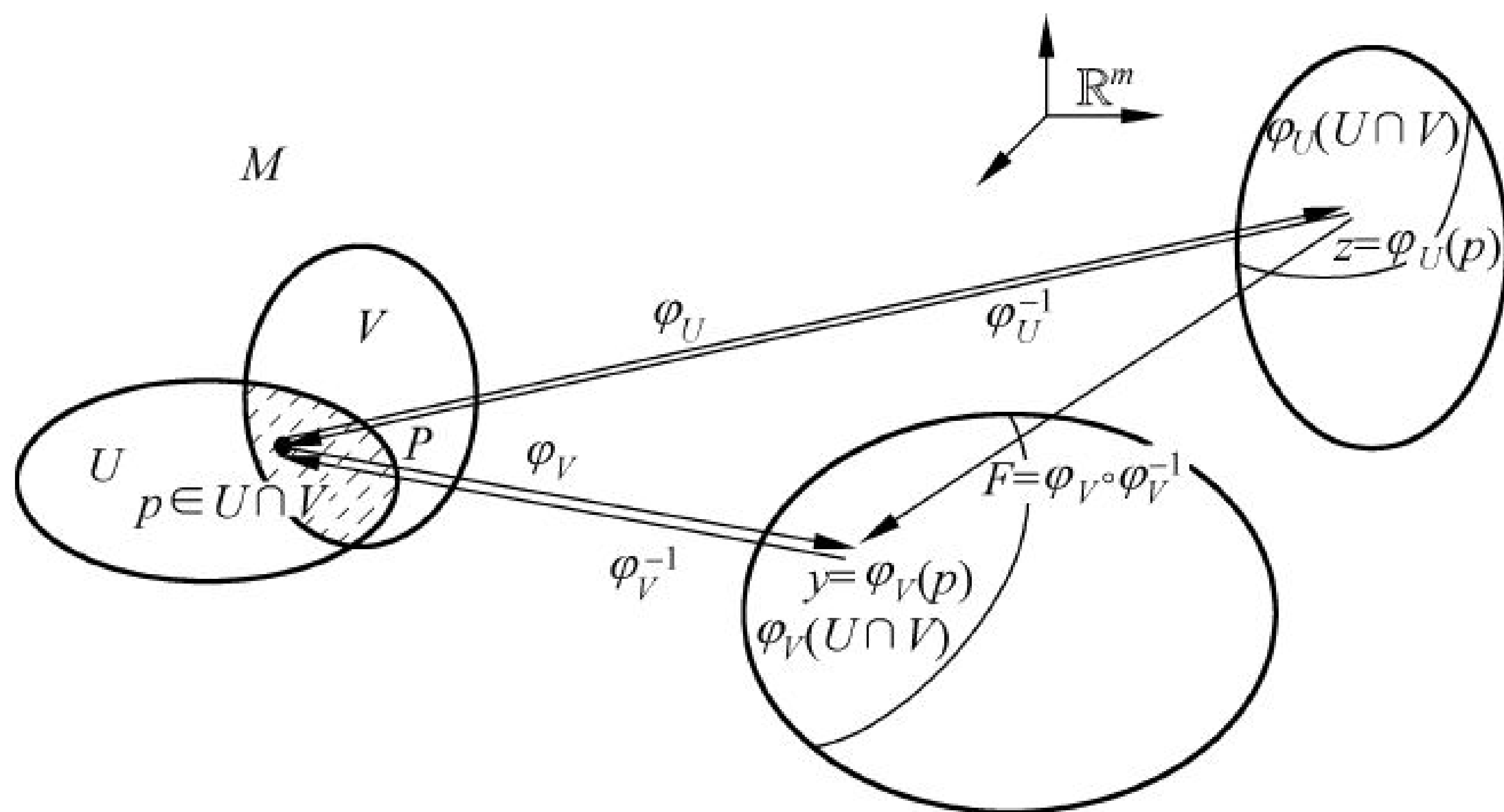


图 6.1.2 $p\in U$ 的局部坐标的相容性

对于 $p\in U$,其局部坐标的相容性如图 6.1.2 所示.且有

$$\begin{aligned}&\forall z\in\varphi_U(U\cap V)\subset\mathbb{R}^m\Rightarrow y\in F(z)\in\varphi_V(U\cap V)\subset\mathbb{R}^m\\&\quad\Rightarrow F(z)=(f_1(z),f_2(z),\cdots,f_m(z))\equiv(y_1,y_2,\cdots,y_m)=y\in\mathbb{R}^m\\&\quad\Rightarrow y=F(z)\Rightarrow y_j=f_j(z)=f_j(z_1,z_2,\cdots,z_m),\quad j=1,2,\cdots,m\\&\quad\Rightarrow z=F^{-1}(y)=(g_1(y),g_2(y),\cdots,g_m(y))\\&\quad\Rightarrow z_k=g_k(y)=g_k(y_1,y_2,\cdots,y_m),\quad k=1,2,\cdots,m\\&\quad\Rightarrow f_j,g_k\text{ 都是连续函数,且有关系}\end{aligned}$$

$$y_j = f_j(z) = f_j(z_1, z_2, \cdots, z_m) = f_j(g_1(y_1, y_2, \cdots, y_m), \cdots, g_m(y_1, y_2, \cdots, y_m)), j = 1, 2, \cdots, m;$$
$$z_j = g_j(y) = g_j(y_1, y_2, \cdots, y_m) = g_j(f_1(z_1, z_2, \cdots, z_m), \cdots, f_m(z_1, z_2, \cdots, z_m)), j = 1, 2, \cdots, m,$$

简记为

$$\begin{cases} y_j = f_j(z_1, \cdots, z_m) = (\varphi_V \circ \varphi_U^{-1}(z_1, \cdots, z_m))_j, \ (z_1, \cdots, z_m) \in \varphi_U(U \cap V), \\ z_j = g_j(y_1, \cdots, y_m) = (\varphi_U \circ \varphi_V^{-1}(y_1, \cdots, y_m))_j, \ (y_1, \cdots, y_m) \in \varphi_V(U \cap V), \end{cases} j = 1, \cdots, m,$$

上述表示可写为

$$\begin{cases} f_j(g_1(y_1, y_2, \cdots, y_m), \cdots, g_m(y_1, y_2, \cdots, y_m)) = y_j, \\ g_j(f_1(z_1, z_2, \cdots, z_m), \cdots, f_m(z_1, z_2, \cdots, z_m)) = z_j, \end{cases} j = 1, 2, \cdots, m. \quad (6.1.1)$$

这样，在一个 T_2 型拓扑空间 (X, τ) 上，就有了局部坐标系

$$\{(u_j = (\varphi_U(p))_j)_{j=1}^m : (U, \varphi_U) \in \mathbb{A}\},$$

这个局部坐标系合理且有意义. 具有满足相容性(1)，(2)的图册 $\mathbb{A} = \{(U, \varphi_U) : U \in \tau\}$ 的 T_2 型拓扑空间 (M, τ) 是一个 m **维拓扑流形**(*m*-**dimensional topological manifold**)，简记为 $(M, \mathbb{A})$，并简称 $(M, \mathbb{A})$ 为 m **维流形**(*m*-**dimensional manifold**).

定义 6.1.3(坐标变换函数、相容条件) 称由(6.1.1)式确定的函数 $f_j(z_1, z_2, \cdots, z_m)$ 与 $g_j(y_1, y_2, \cdots, y_m)$ $(1 \leqslant j \leqslant m)$ 为满足相容性条件的**坐标变换函数**(**coordinate transformation functions**). 当 $(M, \mathbb{A})$ 为 m 维拓扑流形时，坐标变换函数 $f_j(z_1, z_2, \cdots, z_m)$，$g_j(y_1, y_2, \cdots, y_m)$ $(1 \leqslant j \leqslant m)$ 都是连续函数.

若两个坐标卡 (U, φ_U)，(V, φ_V) 的满足相容性条件(6.1.1)的坐标变换函数 $f_j, g_j \in C^r(\mathbb{R}^m)$，$r \in \mathbb{N}$，则称其为 C^r **相容的**(C^r-**compatible**).

定义 6.1.4(C^r 微分流形、光滑流形) 设 $(M, \mathbb{A})$ 是 m 维流形，$\mathbb{A} = \{(U, \varphi_U)\}$ 是 M 上的一个图册，若 $\mathbb{A}$ 满足

(1) $\mathbb{B} = \{U\}$ 是流形 M 的一个**开覆盖**(**open covering**)；

(2) 属于 $\mathbb{A} = \{(U, \varphi_U)\}$ 的任意两个坐标卡是 C^r **相容的**；

(3) $\mathbb{A} = \{(U, \varphi_U)\}$ 是**极大的**，亦即，对于 M 的任一坐标卡 $(\widetilde{U}, \varphi_{\widetilde{U}})$，若它与 $\mathbb{A} = \{(U, \varphi_U)\}$ 中的每一个坐标卡都是 C^r 相容的，必有 $(\widetilde{U}, \varphi_{\widetilde{U}}) \in \mathbb{A}$，

则称 $\mathbb{A} = \{(U, \varphi_U)\}$ 是 M 上的一个 C^r **微分结构**(**differential construction**).

若在 M 上给定一个 C^r 微分结构 $\mathbb{A}$，则称 $(M, \mathbb{A})$ 是一个 C^r **微分流形**(C^r-**differential manifold**)，并称 $\mathbb{A} = \{(U, \varphi_U)\}$ 为微分流形 M 的一个**容许坐标卡系**(**permissible chart system**).

若在 M 上给定一个 C^∞ 微分结构 $\mathbb{A}$，则称 $(M, \mathbb{A})$ 为**光滑流形**(**smooth manifold**).

例 6.1.1 $M = \mathbb{R}^m$.

事实上，取 $U = M$，φ_U 为恒同映射，则 $\mathbb{A} = \{(U, \varphi_U)\}$ 是 $\mathbb{R}^m$ 的一个坐标覆盖，由此确定

$\mathbb{R}^m$ 的光滑流形结构,称为 $\mathbb{R}^m$ 的标准微分结构.

例 6.1.2 在 $\mathbb{R}^2$ 中的单位圆周 $S^1=\{(x_1,x_2)\in\mathbb{R}^2:x_1^2+x_2^2=1\}$ 上构造一个一维光滑流形 M.

解 对于 $S^1=\{(x_1,x_2)\in\mathbb{R}^2:x_1^2+x_2^2=1\}$,记 $x=(x_1,x_2)$,令

$U_1=\{(x_1,x_2)\in S^1:x_2>0\}$, $\varphi_{U_1}(x)=x_1$;

$U_2=\{(x_1,x_2)\in S^1:x_2<0\}$, $\varphi_{U_2}(x)=x_1$;

$U_3=\{(x_1,x_2)\in S^1:x_1>0\}$, $\varphi_{U_3}(x)=x_2$;

$U_4=\{(x_1,x_2)\in S^1:x_1<0\}$, $\varphi_{U_4}(x)=x_2$.

显然 U_1,U_2,U_3,U_4 是 S^1 中的开集,因为 S^1 作为$\mathbb{R}^2$ 的子集,可以考虑子拓扑:将

$$U_1=\{(x_1,x_2)\in S^1:x_2>0\}$$

视为$\mathbb{R}^2$ 中的开集

$$O=\{(x_1,x_2)\in\mathbb{R}^2:x_2>0\}$$

与 S^1 的交,即

$$U_1=\{(x_1,x_2)\in S^1:x_2>0\}=O\cap S^1,$$

U_2,U_3,U_4 类似得到. 则有 $S^1\subset U_1\cup U_2\cup U_3\cup U_4$.

我们证明,$\mathbb{A}=\{(U_j,\varphi_{U_j}):j=1,2,3,4\}$ 给出 S^1 上的一维光滑流形结构,如图 6.1.3 所示.

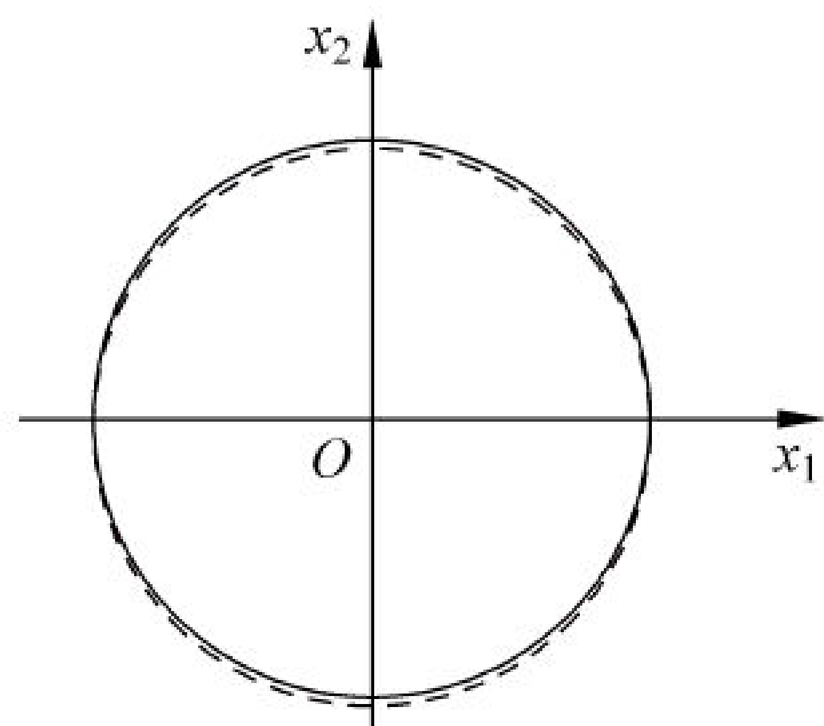


图 6.1.3 S^1 上的一维光滑流形

事实上,$U_1\cap U_2=\varnothing$,$U_3\cap U_4=\varnothing$,故只需考虑 $U_1\cap U_3\neq\varnothing$,$U_1\cap U_4\neq\varnothing$,$U_2\cap U_3\neq\varnothing$,$U_2\cap U_4\neq\varnothing$. 以 $U_1\cap U_4=\{(x_1,x_2)\in S^1:x_1<0,x_2>0\}$ 为例. 因为 $\varphi_{U_1}(x)=x_1$, $\varphi_{U_4}(x)=x_2$,故由

$$\varphi_{U_1}(U_1)=\{z\in\mathbb{R}^1:z=(x_1,0),-1<x_1<1\},$$

$$\varphi_{U_4}(U_4)=\{z\in\mathbb{R}^1:z=(0,x_2),-1<x_2<1\},$$

可得

$$\varphi_{U_1}(U_1\cap U_4)=\{x_1:-1<x_1<0\}\subset\mathbb{R}^1,\quad \varphi_{U_4}(U_1\cap U_4)=\{x_2:0<x_2<1\}\subset\mathbb{R}^1.$$

$\varphi_{U_1}(U_1\cap U_4)$ 与 $\varphi_{U_4}(U_1\cap U_4)$ 均为$\mathbb{R}^1$ 中的非空开集,并且

$$F=\varphi_{U_4}\circ\varphi_{U_1}^{-1}|_{\varphi_{U_1}(U_1\cap U_4)}:\varphi_{U_1}(U_1\cap U_4)\to\varphi_{U_4}(U_1\cap U_4)$$

为同胚映射,使得

$$F=\varphi_{U_4}\circ\varphi_{U_1}^{-1}|_{\varphi_{U_1}(U_1\cap U_4)}:\{x_1:-1<x_1<0\}\subset\mathbb{R}^1\to\{x_2:0<x_2<1\}\subset\mathbb{R}^1,$$

故

$$z=(x_1,0)\in\varphi_{U_1}(U_1\cap U_4)\xrightarrow{F}y=(0,x_2)\in\varphi_{U_4}(U_1\cap U_4). \qquad (*)$$

下面来确定(6.1.1)式中的函数. 记 $z=(z_1,z_2)$,$y=(y_1,y_2)$,则

$$\begin{cases} y_1 = f_1(z_1, z_2), \\ y_2 = f_2(z_1, z_2), \end{cases} \quad \begin{cases} z_1 = g_1(y_1, y_2), \\ z_2 = g_2(y_1, y_2), \end{cases}$$

由(*)式,上两式化为

$$\begin{cases} y_1 = f_1(z_1, z_2) = 0, \\ y_2 = f_2(z_1, z_2) = x_2, \end{cases} \quad \begin{cases} z_1 = g_1(y_1, y_2) = x_1, \\ z_2 = g_2(y_1, y_2) = 0; \end{cases}$$

因此

$$x_2 = f_2(z_1, z_2) = f_2(x_1, 0) \equiv f_2(x_1) = \sqrt{1 - x_1^2}, \quad -1 < x_1 < 0;$$

$$x_1 = g_1(y_1, y_2) = g_1(0, x_2) \equiv g_1(x_2) = -\sqrt{1 - x_2^2}, \quad 0 < x_2 < 1.$$

易见,两个坐标卡(U_1, φ_{U_1}),(U_4, φ_{U_4})满足相容条件(6.1.1),且坐标变换函数

$$x_2 = f_2(x_1) = \sqrt{1 - x_1^2}, \quad -1 < x_1 < 0$$

与

$$x_1 = g_1(x_2) = -\sqrt{1 - x_2^2}, \quad 0 < x_2 < 1$$

都是 C^∞ 函数,故为 C^∞ 相容的.

其余三种情况 $U_1 \cap U_3 \neq \varnothing$,$U_2 \cap U_3 \neq \varnothing$,$U_2 \cap U_4 \neq \varnothing$ 可类似讨论.

由以上证明,可知 $S^1 = \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 = 1\}$ 上赋予 $\mathbb{A} = \{U_1, U_2, U_3, U_4\}$,就使得 S^1 成为 $\mathbb{R}^2$ 的一维光滑流形 $M = S^1$.

注 在同一个拓扑流形上,可以有不同的微分结构. J. Milnor 在 1956 年给出著名的"Milnor 怪球"的例子,说明在同胚的拓扑流形上存在不同构的微分流形结构,因此微分结构有独立于拓扑结构的意义.

2. 微分流形上映射的光滑性

1) 光滑流形上的光滑函数

定义 6.1.5(光滑流形上的光滑函数) 设$(M, \mathbb{A})$是 m 维光滑流形. $f: M \to \mathbb{R}$ 是(M, τ)上的实值函数,对于 $p \in M$,若$(U, \varphi_U) \in \mathbb{A}$ 是包含点 p 的容许坐标卡,则复合函数 $f \circ \varphi_U^{-1}$ 是开集 $\varphi_U(U) \subset \mathbb{R}^m$ 上的实值函数 $f \circ \varphi_U^{-1} : \varphi_U(U) \to \mathbb{R}$,如图 6.1.4 所示.

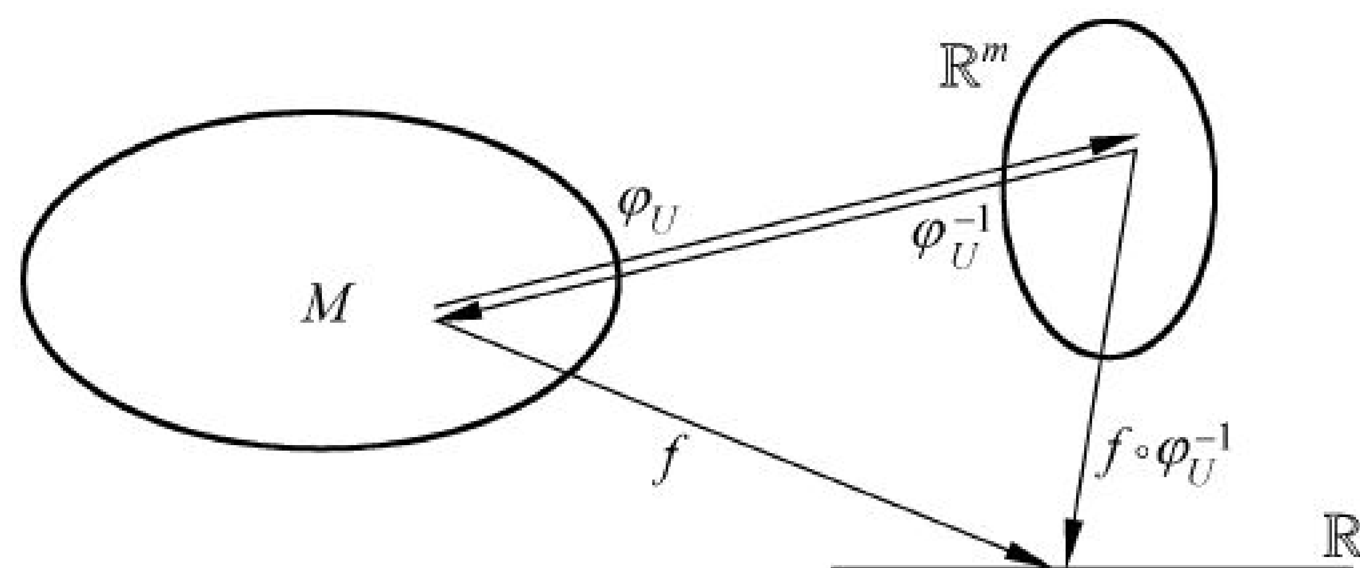


图 6.1.4 流形上的光滑函数

于是,如果 $\varphi_U(U)\subset\mathbb{R}^m$ 上的实值函数

$$f\circ\varphi_U^{-1}:\varphi_U(U)\to\mathbb{R}$$

在 $\varphi_U(U)\subset\mathbb{R}^m$ 中的点 $\varphi_U(p)\in\varphi_U(U)$ 处是 C^∞ 的,则称**流形 $(M,\mathbb{A})$ 上的实值函数 $f:M\to\mathbb{R}$ 在点 $\varphi_U(p)$ 是 C^∞ 的**. 若 $f:M\to\mathbb{R}$ 在 $(M,\mathbb{A})$ 的每一点都是 C^∞ 的,则称**流形 $(M,\mathbb{A})$ 上的函数 $f:M\to\mathbb{R}$ 在 $(M,\mathbb{A})$ 上是 C^∞ 的.**

记 M 上的 C^∞ 函数的全体为 $C^\infty(M)=\{f:f$ 是 M 上的无穷次可微的连续函数$\}$. 于是,**流形 $(M,\mathbb{A})$ 上的光滑函数**可表述为

$f:M\to\mathbb{R}$ **是流形 $(M,\mathbb{A})$ 上光滑函数** $\Leftrightarrow$ $f\circ\varphi_U^{-1}:\mathbb{R}^m\to\mathbb{R}$ **为光滑函数**

2) 光滑流形之间的光滑映射

定义 6.1.6(光滑流形之间的光滑映射) 设 $(M,\mathbb{A})$ 与 $(N,\mathbb{B})$ 分别是 m 维与 n 维光滑流形,$f:M\to N$ 是 (M,τ) 到 (N,υ) 的连续映射. 若对于点 $p\in M$ 与点 $f(p)\in N$,分别存在坐标卡 $(U,\varphi_U)\in\mathbb{A}$ 与 $(V,\psi_V)\in\mathbb{B}$,使得映射

$$\psi_V\circ f\circ\varphi_U^{-1}:\varphi_U(U)\subset\mathbb{R}^m\to\psi_V(V)\subset\mathbb{R}^n$$

在点 $\varphi_U(p)\in\varphi_U(U)$ 是 C^∞ 的,则称**映射** $f:M\to N$ 在点 $p\in M$ 是 C^∞ 的;若在 M 的每一点都是 C^∞ 的,则称 $f:M\to N$ 在 M 上是 C^∞ 的,或称 $f:M\to N$ 是流形 $(M,\mathbb{A})$ 到流形 $(N,\mathbb{B})$ **的光滑映射.**

注 (1)光滑函数是光滑流形之间的光滑映射的特例;(2)映射的光滑性与光滑流形的容许坐标卡的选取无关.

若 $\dim M=\dim N$,且 $f:M\to N$ 是拓扑同胚,当 f 与 f^{-1} 都是光滑映射时,则称 $f:M\to N$ 是**可微同胚(differentiable homeomorphism)**,并称两流形 $(M,\mathbb{A})$ 与 $(N,\mathbb{B})$ **可微同胚**,也称 $(M,\mathbb{A})$ 与 $(N,\mathbb{B})$ 的光滑流形结构是**同构的(isomorphic)**.

于是,流形 $(M,\mathbb{A})$ 到流形 $(N,\mathbb{B})$ 的光滑映射可表述为

$f:M\to N$ 是流形 $(M,\mathbb{A})$ 到流形 $(N,\mathbb{B})$ 的光滑映射

$\Leftrightarrow$ $\psi_V\circ f\circ\varphi_U^{-1}:\varphi_U(U)\to\psi_V(V)$ 是 $\mathbb{R}^m$ 到 $\mathbb{R}^n$ 的光滑映射

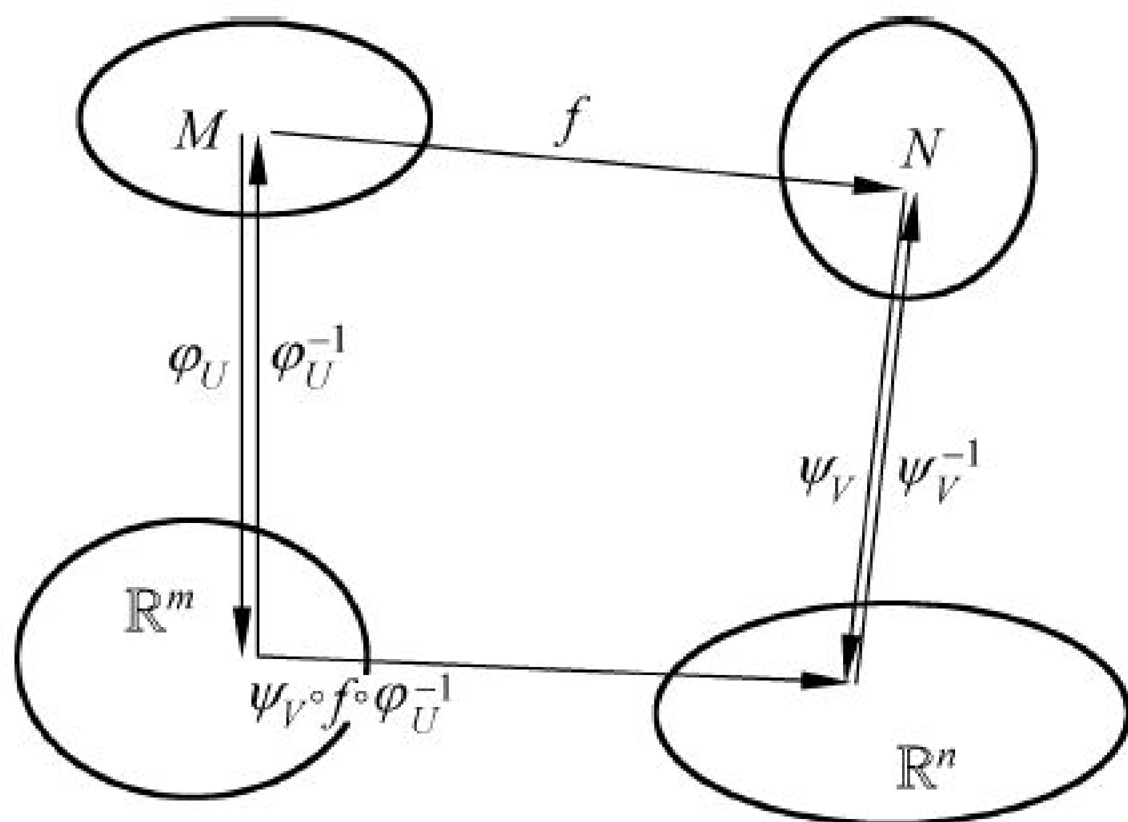


图 6.1.5 流形之间的光滑映射

3）光滑流形上的光滑参数曲线

定义 6.1.7（光滑流形上的光滑参数曲线） 取$(M,\tau)=((a,b),\tau)$，τ 为$\mathbb{R}^1$上的通常拓扑，$(N,\mathbb{B})$是 n 维光滑流形，若映射 $f:(a,b)\to N$ 是(a,b)到 N 的光滑映射，则称映射 $f:(a,b)\to N$ 为流形$(N,\mathbb{B})$上的**光滑（参数）曲线（smooth（parameter）curve）**，如图 6.1.6 所示.

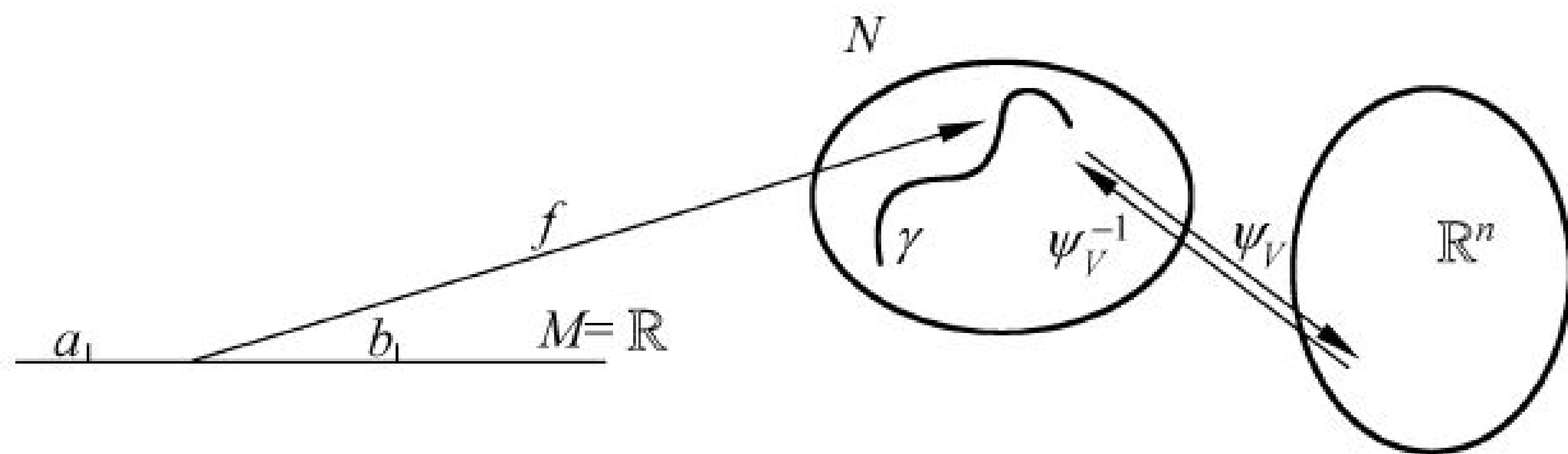


图 6.1.6 流形上的光滑参数曲线

事实上，视$(M,\tau)=((a,b),\tau)$，τ 为$\mathbb{R}^1$上的拓扑，亦即取$(M,\mathbb{A})$为一维流形；设$(N,\mathbb{B})$是 n 维光滑流形，N 的坐标卡是$\mathbb{B}=\{(V,\psi_V)\}$. 视映射 $f:M\to N$ 的定义域为 $\mathfrak{D}=(a,b)\subset M$，值域为 $\mathfrak{R}_f\subset N$. 若映射 $f:(a,b)\to N$ 是 $\mathfrak{D}\subset M$ 到 N 的光滑映射，则映射 $f:(a,b)\to N$ 为 N 上的一条光滑曲线.

流形$(N,\mathbb{B})$上的光滑曲线可表述为

$(N,\mathbb{B})$上的曲线 $f:(a,b)\to N$ 为光滑曲线 $\Leftrightarrow$ $\psi_V\circ f:(a,b)\to\mathbb{R}^n$ 是光滑映射

$$\updownarrow \qquad\qquad\qquad\qquad \updownarrow$$

$f:\mathfrak{D}_f=(a,b)\subset M\to\mathfrak{R}_f=N$ $\qquad$ $g=\psi_V\circ f:(a,b)\to\mathbb{R}^n$

↑ 一维流形M到n维流形N的映射 $\qquad$ ↑ 一元n维向量值函数

借助于局部坐标系，可以定义流形上的函数、流形之间映射以及流形上的曲线，并且定义它们的光滑性.

注 流形上的参数曲线是流形间光滑映射另一个重要特例. 今后我们对 $f:(a,b)\to N$ 与光滑曲线不加区别，统一记为 $\gamma:(a,b)\to N$.

6.1.2 余切空间、切空间

三维欧氏空间的几何直观告诉我们：光滑曲线在每一点有切线，光滑曲面在每一点有切平面. 切平面可以定义为光滑曲面上过一个点的每一条光滑曲线的切线都在同一平面时，这个平面就是该点的切平面. 在一个光滑流形$(M,\mathbb{A})$，$\mathbb{A}=\{(U,\varphi_U)\}$上，当如何推广这些概念呢？本节来解决这个问题.

给定一个 m 维光滑流形$(M,\mathbb{A})$，$\mathbb{A}=\{(U,\varphi_U)\}$.

1. 基本空间 $C_p^\infty(M)$

集合

$$C_p^\infty(M)=\{f:U\to\mathbb{R},f\in C^\infty(U),\ p\in U,U\in\tau\}$$

表示由 $p\in M$ 的一个邻域 $U\equiv U_p$ 到$\mathbb{R}$ 的 C^∞ 实值函数 $f:U\to\mathbb{R}$ 的全体. f 的定义域就是 U，用通常的记号，$\mathfrak{D}_f=U$.

设 $f,g\in C_p^\infty(M)$，且 $\mathfrak{D}_f=U$，$\mathfrak{D}_g=V$，在集合 $C_p^\infty(M)$上定义运算如下：

加法 $(f+g)(x)=f(x)+g(x)$，$\forall x\in\mathfrak{D}_f\cap\mathfrak{D}_g$；

数乘 $(\alpha f)(x)=\alpha f(x)$，$\forall x\in\mathfrak{D}_f$，$\alpha\in\mathbb{R}$；

于是，$(C_p^\infty(M),+,\alpha\cdot)$成为实数域$\mathbb{R}$ 上的一个实线性空间；进而，还可定义

乘法 $(f\cdot g)(x)=f(x)g(x)$，$\forall x\in\mathfrak{D}_f\cap\mathfrak{D}_g$.

注意，$C_p^\infty(M)$中的函数 $f:U\to\mathbb{R}$ 的 C^∞ 光滑性是在定义 6.1.5 意义下的.

2. 函数芽空间 $\mathfrak{F}_p$

① $C_p^\infty(M)$中的等价关系～　对于 $f,g\in C_p^\infty(M)$，若∃点 $p\in M$ 的邻域 H，$p\in H\subset M$，使得 $f|_H=g|_H$，**则称 f 与 g 等价**，记为 $f\sim g$.

记与 $f\in C_p^\infty(M)$等价的、$C_p^\infty(M)$中的函数的全体为$[f]$，称**等价函数的集合**$[f]=\{g\in C_p^\infty(M):g\sim f\}$为**流形 M 在点 $p\in M$ 的 C^∞ 函数芽(germ)**.

空间 $C_p^\infty(M)$等价关系～的意义：把只要在 p 点的某一个邻域 H 中相等的所有函数视为“一个”，若 $f,g\in C_p^\infty(M)$，且 $f|_H=g|_H$，则称为二者“等价”：$f\sim g$.

② 等价类空间　定义商集 $\mathfrak{F}_p=C_p^\infty(M)/\sim\equiv\{[f]:f\in C_p^\infty(M)\}$，$\mathfrak{F}_p$ 为 C^∞ 函数芽集.

定义函数芽集 $\mathfrak{F}_p=C_p^\infty(M)/\sim$ 中的运算：

加法 $[f],[g]\in\mathfrak{F}_p\Rightarrow[f]+[g]=[f+g]$；

数乘 $[f]\in\mathfrak{F}_p$，$\alpha\in\mathbb{R}\Rightarrow\alpha[f]=[\alpha f]$，

于是，函数芽集 $\mathfrak{F}_p=C_p^\infty(M)/\sim$ 成为实线性空间，称为**函数芽空间(germ space)**. 显然，它不是一个有限维空间.

函数芽空间 $\mathfrak{F}_p=C_p^\infty(M)/\sim$的意义：$\mathfrak{F}_p$ 空间中的元$[f]$都是等价类，称为函数芽，每个芽$[f]\in\mathfrak{F}_p$ 中有无穷多个函数 $g\in[f]$，它们在 $p\in M$ 的一个邻域 H 中，与 f 有相同的值. 这样无穷多个 g 中，只要抽出一个 f 作为代表，相当于出一个芽$[f]$. 如此划分 $C_p^\infty(M)$中的等价类，提供了按等价类研究函数性质的一个重要方法.

3. 参数曲线空间 Γ_p

流形$(M,\mathbb{A})$上过点 $p\in M$ 的参数曲线 $\gamma:(-\delta,\delta)\to M$　对于点 $p\in M$，若存在 $\delta>0$，使得 $0\leftrightarrow p$，且 $\gamma:(-\delta,\delta)\to M$ 是一对一映射，则称映射 $\gamma:(-\delta,\delta)\to M$ 为流形$(M,\mathbb{A})$上过 $p\in M$ 点的参数曲线. 记$(M,\mathbb{A})$上过 $p\in M$ 点的光滑参数曲线的全体为

$$\Gamma_p=\{\gamma:(-\delta,\delta)\to M,\gamma\text{ 是 }C^\infty\text{ 光滑的}\},$$

称 Γ_p 为**过 p 点的光滑参数曲线集**，简称参数曲线集.

在 Γ_p 中定义运算

加法 $(\gamma_1+\gamma_2)(t)=\gamma_1(t)+\gamma_2(t)$，$\forall t\in(-\delta,\delta)$；

数乘 $(\alpha\gamma)(t)=\alpha\gamma(t)$，$\forall t\in(-\delta,\delta)$，$\alpha\in\mathbb{R}$，

则 Γ_p 成为实线性空间，称为**参数曲线空间**.

参数曲线空间 Γ_p 的意义：过点 $p\in M$ 有无穷多条光滑参数曲线 γ，将它们的参数限制在$t\in(-\delta,\delta)$中(须存在一个共同的 $\delta>0$)，并且使得参数 $t=0$ 对应于点 p，记为 $\gamma(0)=p$.

4. 零配合子空间 $\mathfrak{I}_p$

函数芽$[f]\in\mathfrak{F}_p$ 与参数曲线 $\gamma\in\Gamma_p$ 的配合 定义$[f]\in\mathfrak{F}_p$ 与 $\gamma\in\Gamma_p$的配合$\langle\langle\gamma,[f]\rangle\rangle$：对于$[f]\in\mathfrak{F}_p$与 $\gamma\in\Gamma_p$，称

$$\langle\langle\gamma,[f]\rangle\rangle=\left.\frac{\mathrm{d}(f\circ\gamma)}{\mathrm{d}t}\right|_{t=0},\quad t\in(-\delta,\delta) \tag{6.1.2}$$

为$[f]\in\mathfrak{F}_p$ 与 $\gamma\in\Gamma_p$ 的配合，如图 6.1.7 所示. 显然，$f\circ\gamma:(-\delta,\delta)\to\mathbb{R}$ 的光滑性有意义.

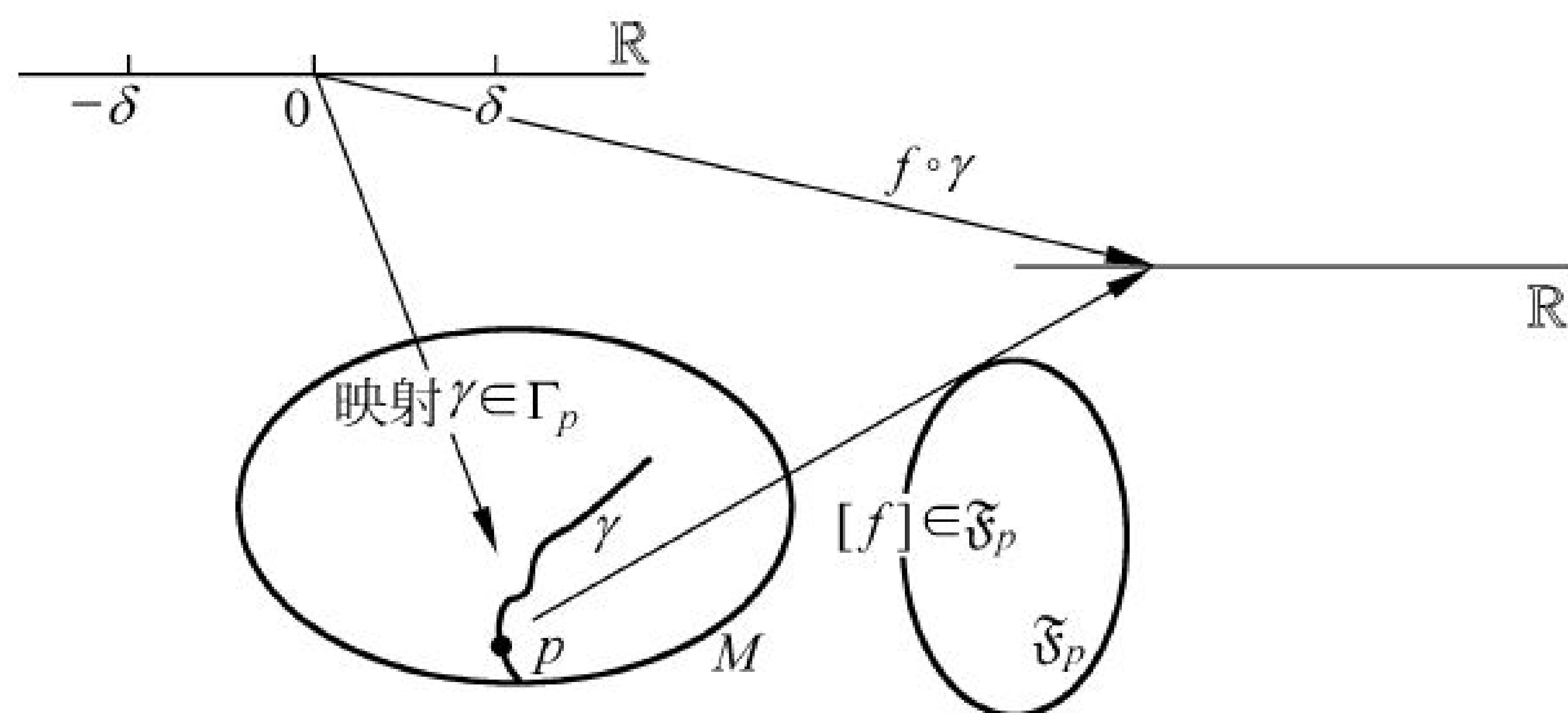


图 6.1.7 γ 与$[f]$的配合

配合的意义：$\langle\langle\gamma,[f]\rangle\rangle=\left.\dfrac{\mathrm{d}(f\circ\gamma)}{\mathrm{d}t}\right|_{t=0}$ 相当于导数$\dfrac{\mathrm{d}(f\circ\gamma)}{\mathrm{d}t}$ 在 $t=0$ 的值，亦即，它是映射 $f\circ\gamma$ 在点 $p(t=0)$的切矢量.

注意，配合$\langle\langle\gamma,[f]\rangle\rangle$与 f 的选择无关(为什么?). 于是，可以定义函数芽空间 $\mathfrak{F}_p$ 的子集

$$\mathfrak{I}_p=\{[f]\in\mathfrak{F}_p:\langle\langle\gamma,[f]\rangle\rangle=0,\ \forall\,\gamma\in\Gamma_p\}\subset\mathfrak{F}_p,$$

称 $\mathfrak{I}_p$ 为零配合集.

在零配合集 $\mathfrak{I}_p\subset\mathfrak{F}_p$ 中，定义

$$[f],[g]\in\mathfrak{I}_p\Rightarrow\forall\,\gamma\in\Gamma_p,\ [f]+[g]=[f+g]; \tag{加法}$$

$$[f]\in\mathfrak{I}_p,\alpha\in\mathbb{R}\Rightarrow\forall\,\gamma\in\Gamma_p,\ \alpha[f]=[\alpha f]. \tag{数乘}$$

则上述运算封闭. 这是因为

$$\begin{aligned}
[f],[g]\in\mathfrak{I}_p\Rightarrow\forall\,\gamma\in\Gamma_p,\text{有}\langle\langle\gamma,[f+g]\rangle\rangle&=\left.\frac{\mathrm{d}(f+g)\circ\gamma}{\mathrm{d}x}\right|_{t=0}\\
&=\left.\frac{\mathrm{d}(f\circ\gamma)}{\mathrm{d}x}\right|_{t=0}+\left.\frac{\mathrm{d}(g\circ\gamma)}{\mathrm{d}x}\right|_{t=0}\\
&=\langle\langle\gamma,[f]\rangle\rangle+\langle\langle\gamma,[g]\rangle\rangle=0
\end{aligned}$$

$$\Rightarrow\langle\langle\gamma,[f+g]\rangle\rangle=0\Rightarrow[f+g]\in\mathfrak{I}_p;$$

$$[f]\in\mathfrak{J}_p,\ \alpha\in\mathbb{R}\Rightarrow\forall\gamma\in\Gamma_p,\text{有}\langle\langle\gamma,\alpha[f]\rangle\rangle=\left.\frac{\mathrm{d}(\alpha f\circ\gamma)}{\mathrm{d}t}\right|_{t=0}=\alpha\left.\frac{\mathrm{d}(f\circ\gamma)}{\mathrm{d}t}\right|_{t=0}$$
$$=\alpha\langle\langle\gamma,[f]\rangle\rangle=0$$
$$\Rightarrow\alpha[f]\in\mathfrak{J}_p\Rightarrow\langle\langle\gamma,\alpha[f]\rangle\rangle=0\Rightarrow\alpha[f]\in\mathfrak{J}_p.$$

于是

$$\mathfrak{J}_p=\{[f]\in\mathfrak{F}_p:\langle\langle\gamma,[f]\rangle\rangle=0,\forall\gamma\in\Gamma_p\}$$

是 $\mathfrak{F}_p$ 的实线性子空间,称 $\mathfrak{J}_p$ 为 $\mathfrak{F}_p$ 的**零配合子空间**.

零配合子空间有表征定理.

定理 6.1.1 设$(M,\mathbb{A})$为 m 维光滑流形,$\mathfrak{F}_p$ 为 C^∞ 函数芽空间,$\mathfrak{J}_p$ 为零配合子空间. 若$[f]\in\mathfrak{F}_p$,则$[f]\in\mathfrak{J}_p$当且仅当对于包含 $p\in M$ 点的容许坐标卡(U,φ_U),有

$$\left.\frac{\partial(f\circ\varphi_U^{-1})}{\partial x_j}\right|_{\varphi_U(p)}=0,\quad 1\leqslant j\leqslant m.$$

证 设(U,φ_U)为包含 $p\in M$ 的容许坐标卡,于是,对于$[f]\in\mathfrak{F}_p$,令

$$f\circ\varphi_U^{-1}(x_1,x_2,\cdots,x_m)\equiv F(x_1,x_2,\cdots,x_m),\quad (x_1,x_2,\cdots,x_m)\in\varphi_U(U).$$

另一方面,对于 $\gamma\in\Gamma_p$,$\gamma:(-\delta,\delta)\to U\subset M$,记其关于容许坐标卡$(U,\varphi_U)$的坐标表示为

$$((\varphi_U\circ\gamma(t))_1,\cdots,(\varphi_U\circ\gamma(t))_m)=(x_1(t),\cdots,x_m(t))\in\mathbb{R}^m,\ t\in(-\delta,\delta).$$

于是,对此$[f]\in\mathfrak{F}_p$ 与 $\gamma\in\Gamma_p$,做配合

$$\langle\langle\gamma,[f]\rangle\rangle=\left.\frac{\mathrm{d}(f\circ\gamma)}{\mathrm{d}t}\right|_{t=0}=\left.\frac{\mathrm{d}(f\circ\varphi_U^{-1}\circ\varphi_U\circ\gamma)}{\mathrm{d}t}\right|_{t=0}=\left.\frac{\mathrm{d}(F\circ\varphi_U\circ\gamma)}{\mathrm{d}t}\right|_{t=0}$$
$$=\left.\frac{\mathrm{d}(F(x_1(t),\cdots,x_m(t)))}{\mathrm{d}t}\right|_{t=0}=\sum_{j=1}^m\left.\frac{\partial F}{\partial x_j}\right|_{\varphi_U(p)}\left.\frac{\mathrm{d}x_j(t)}{\mathrm{d}t}\right|_{t=0},$$

得

$$[f]\in\mathfrak{J}_p\Leftrightarrow\forall\gamma\in\Gamma_p,\langle\langle\gamma,[f]\rangle\rangle=0\Leftrightarrow\forall\gamma\in\Gamma_p,\sum_{j=1}^m\left.\frac{\partial F}{\partial x_j}\right|_{\varphi_U(p)}\left.\frac{\mathrm{d}x_j(t)}{\mathrm{d}t}\right|_{t=0}=0$$
$$\Leftrightarrow\left.\frac{\partial F}{\partial x_j}\right|_{\varphi_U(p)}=0,1\leqslant j\leqslant m\Leftrightarrow\left.\frac{\partial(f\circ\varphi_U^{-1})}{\partial x_j}\right|_{\varphi_U(p)}=0,\ 1\leqslant j\leqslant m.$$

定理得证.

零配合子空间 $\mathfrak{J}_p=\{[f]\in\mathfrak{F}_p:\langle\langle\gamma,[f]\rangle\rangle=0,\forall\gamma\in\Gamma_p\}$**的意义**:零配合子空间 $\mathfrak{J}_p$ 中的元$[f]$,恰好是在点 $p\in M$ 关于局部坐标的偏导数的值都是零的光滑函数芽所构成的线性子空间. 当 γ 与 γ'等价时,$\gamma\sim\gamma'$蕴含一元函数 $f\circ\gamma$ 与 $f\circ\gamma'$在 $t=0$ 有相同导数.

5. 余切空间 $T_p^*=\mathfrak{F}_p/\sim$——芽空间 $\mathfrak{F}_p$ 中的等价类

定义 6.1.8(余切空间) 设 $\mathfrak{F}_p$ 是函数芽空间,$\mathfrak{J}_p$ 是其零配合子空间,定义商集 $T_p^*=\mathfrak{F}_p/\sim$;亦即,定义芽空间 $\mathfrak{F}_p$ 中的元的等价关系$\sim$:对于$[f],[g]\in\mathfrak{F}_p$,

$$[f]\sim[g]\Leftrightarrow[f]-[g]\in\mathfrak{J}_p.$$

在上述等价关系之下，记与$[f]\in\mathfrak{F}$等价的元$[g]\in\mathfrak{F}$全体所成的集合为

$$[f]^*=\{[g]\in\mathfrak{F}:[g]\sim[f]\},$$

因此，与$[f]\in\mathfrak{F}$等价的$[g]\in[f]^*$，满足：$\forall\gamma\in\Gamma_p$，

$$[g]\in[f]^*\Leftrightarrow\langle\langle\gamma,[f]-[g]\rangle\rangle=0\Leftrightarrow\langle\langle\gamma,[f]\rangle\rangle=\langle\langle\gamma,[g]\rangle\rangle,$$

亦即$[f],[g]\in\mathfrak{F}$，$[f]\sim[g]$表示$[f],[g]$与任一$\gamma\in\Gamma_p$的配合相等.

商集$T_p^*=\mathfrak{F}/\sim=\{[f]^*:[f]\in\mathfrak{F}\}$的运算结构：在商集$T_p^*=\mathfrak{F}/\sim$上定义

加法　　$[f]^*,[g]^*\in T_p^*\Rightarrow[f]^*+[g]^*=[f+g]^*$；

数乘　　$[f]^*\in T_p^*,\alpha\in\mathbb{R}\Rightarrow\alpha[f]^*=[\alpha f]^*$，

则$T_p^*=\mathfrak{F}/\sim$成为实线性空间，称为流形M在点$p\in M$的**余切空间(cotangent space)**. 余切空间$T_p^*=\mathfrak{F}/\sim$中的元$[f]^*\in T_p^*$也记为$[f]^*=(\mathrm{d}f)_p$，称为流形M在点$p\in M$的**余切矢量(cotangent vector)**.

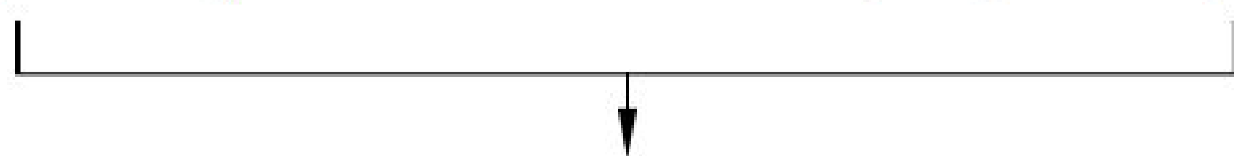
6. 余切矢量$(\mathrm{d}f)_p$的一个公式

对于流形M上任意s个光滑函数$f^1,f^2,\cdots,f^s\in C_p^\infty(M)$，以及$\mathbb{R}^s$上的一个光滑函数

$$F(y^1,y^2,\cdots,y^s):\mathbb{R}^s\to\mathbb{R},$$

则$F(f^1,f^2,\cdots,f^s):\mathbb{R}^s\to\mathbb{R}$是在点$(f^1(p),f^2(p),\cdots,f^s(p))\in\mathbb{R}^s$的邻域内的光滑函数. 由此，得到满足$f(q)\equiv F(f^1(q),f^1(q),\cdots,f^s(q))$，$q\in U_p$的复合函数$f:M\to\mathbb{R}$，

$$f^1,f^2,\cdots,f^s\in C_p^\infty(M)\qquad\qquad F(f^1,f^2,\cdots,f^s):\mathbb{R}^s\to\mathbb{R}$$



$$f(q)\equiv F(f^1(q),f^2(q),\cdots,f^s(q)),\quad q\in U_p$$

对于流形M上这样的函数$f:M\to\mathbb{R}$的余切矢量$(\mathrm{d}f)_p=[f]^*$有如下表示公式.

定理 6.1.2　设M为m维光滑流形，对于任意的$f^1,f^2,\cdots,f^s\in C_p^\infty(M)$，对于$\mathbb{R}^s$上的光滑函数$F(y^1,y^2,\cdots,y^s):\mathbb{R}^s\to\mathbb{R}$，令

$$f(q)\equiv F(f^1(q),f^2(q),\cdots,f^s(q)),\quad q\in U_p,$$

则$f\in C_p^\infty(M)$，且成立公式

$$(\mathrm{d}f)_p=\sum_{k=1}^{s}\left(\frac{\partial F}{\partial f^k}\right)_{(f^1(p),\cdots,f^s(p))}\cdot(\mathrm{d}f^k)_p.\qquad(6.1.3)$$

证　设每个$f^1,f^2,\cdots,f^s\in C_p^\infty(M)$的定义域为$U_k(p\in U_k,1\leqslant k\leqslant s)$，则$f$的定义域为

$$U\equiv\bigcap_{k=1}^{s}U_k.$$

于是，对于$q\in U\equiv\bigcap_{k=1}^{s}U_k$，令

$$f(q)\equiv F(f^1(q),f^2(q),\cdots,f^s(q)),\quad q\in U_p,$$

由 $F\in C^{\infty}(\mathbb{R}^s)$，故 $f\in C_p^{\infty}(M)$. 对 $\forall\gamma\in\Gamma_p$，计算

$$\begin{aligned}\langle\langle\gamma,[f]\rangle\rangle &= \left.\frac{\mathrm{d}(f\circ\gamma)}{\mathrm{d}t}\right|_{t=0} = \left.\frac{\mathrm{d}F(f^1\circ\gamma(t),\cdots,f^s\circ\gamma(t))}{\mathrm{d}t}\right|_{t=0}\\ &= \sum_{k=1}^{s}\left(\frac{\partial F}{\partial f^k}\right)_{(f^1(p),\cdots,f^s(p))}\left.\frac{\mathrm{d}(f^k\circ\gamma(t))}{\mathrm{d}t}\right|_{t=0}\\ &= \sum_{k=1}^{s}\left(\frac{\partial F}{\partial f^k}\right)_{(f^1(p),\cdots,f^s(p))}\langle\langle\gamma,[f^k]\rangle\rangle\\ &= \left\langle\left\langle\gamma,\sum_{k=1}^{s}\left(\frac{\partial F}{\partial f^k}\right)_{(f^1(p),\cdots,f^s(p))}[f^k]\right\rangle\right\rangle.\end{aligned}$$

因此

$$\langle\langle\gamma,[f]\rangle\rangle = \left\langle\left\langle\gamma,\sum_{k=1}^{s}\left(\frac{\partial F}{\partial f^k}\right)_{(f^1(p),\cdots,f^s(p))}[f^k]\right\rangle\right\rangle,$$

上式蕴含

$$[f]-\sum_{k=1}^{s}\left(\frac{\partial F}{\partial f^k}\right)_{(f^1(p),\cdots,f^s(p))}[f^k]\in\mathfrak{J}_p,$$

亦即

$$(\mathrm{d}f)_p = \sum_{k=1}^{s}\left(\frac{\partial F}{\partial f^k}\right)_{(f^1(p),\cdots,f^s(p))}(\mathrm{d}f^k)_p.$$

定理得证.

图 6.1.8 给出了定理 6.1.2 的图形表示.

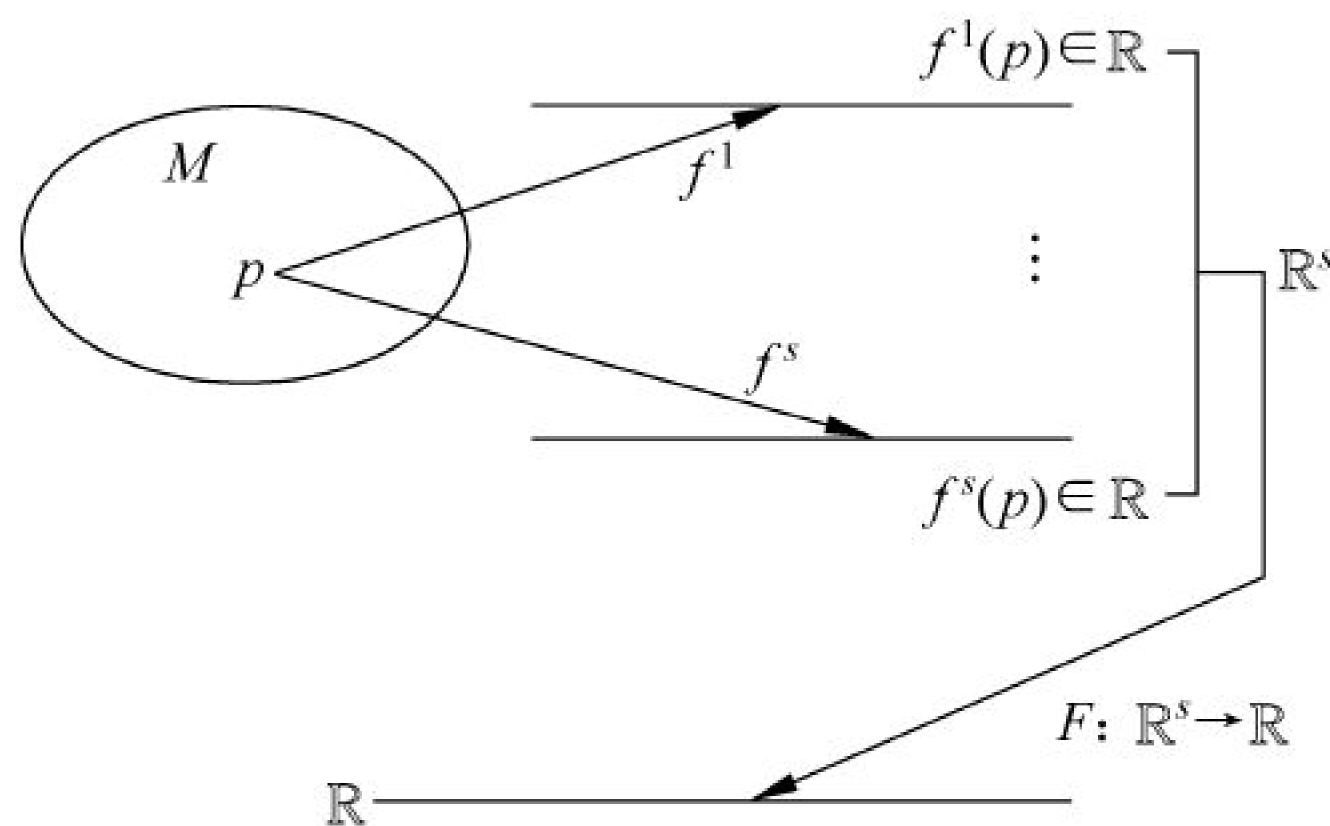


图 6.1.8 定理 6.1.2 的图解

由以上定理可得到余切矢量 $(\mathrm{d}f)_p=[f]^*$ 的重要的性质.

定理 6.1.3 设 $f,g\in C_p^{\infty}(M)$，$\alpha,\beta\in\mathbb{R}$，有 $(\mathrm{d}f)_p=[f]^*\in T_p^*$，$(\mathrm{d}g)_p=[g]^*\in T_p^*$，则

$$(\mathrm{d}(\alpha f+\beta g))_p = \alpha(\mathrm{d}f)_p+\beta(\mathrm{d}g)_p;$$
$$(\mathrm{d}(f\cdot g))_p = (\mathrm{d}f)_p\,g(p)+(\mathrm{d}g)_p\,f(p).$$

这一定理的证明由微积分中的基本公式得到，因为 $(\mathrm{d}f)_p$ 运算实际上已经化为微积分

中的求导运算.

定理 6.1.4(余切空间 T_p^* 的维数定理) 设$(M,\mathbb{A})$为 m 维光滑流形,则 $\dim T_p^*=m$.

证 取点 $p\in M$ 的一个容许坐标卡为(U,φ_U),$\forall q\in U$,其局部坐标$(u_1,u_2,\cdots,u_m)$记为

$$\begin{aligned}(u_1(q),\cdots,u_m(q))&=((\varphi_U(q))_1,\cdots,(\varphi_U(q))_m)\\&=((e_1\circ\varphi_U)(q),\cdots,(e_m\circ\varphi_U)(q)),\end{aligned}\qquad(6.1.4)$$

其中$\{e_1,\cdots,e_m\}$为$\mathbb{R}^m$ 中选定的一个坐标系,而 $\forall q\in U\subset M\Rightarrow\varphi_U(q)\in\mathbb{R}^m$.

图 6.1.9 给出了定理 6.1.4 的图形表示.

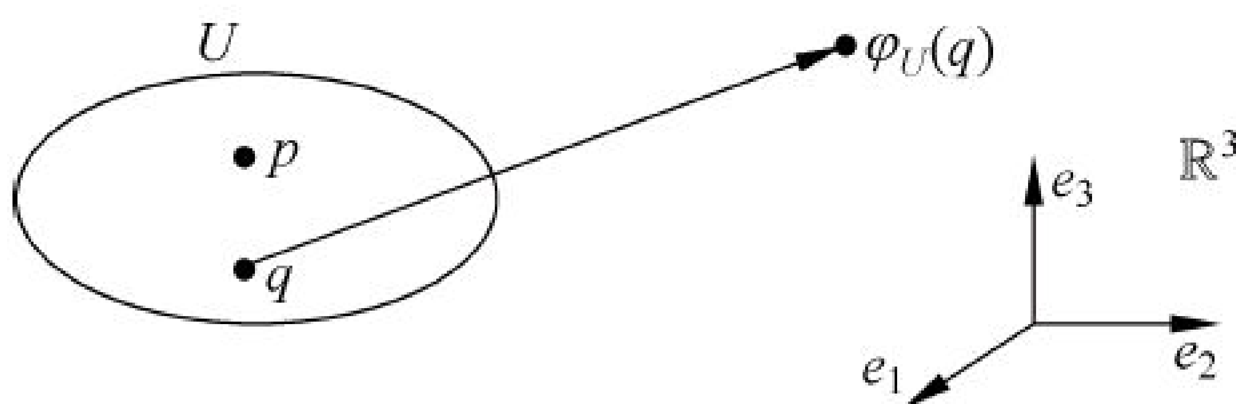


图 6.1.9 定理 6.1.4 的图解

例如,$m=3$ 时,有 $\varphi_U(q)=(\varphi_U(q))_1e_1+(\varphi_U(q))_2e_2+(\varphi_U(q))_3e_3$,此式右边可视为

$$q\in U\Rightarrow\varphi_U(q)\in\mathbb{R}^3\Rightarrow((\varphi_U(q))_1,(\varphi_U(q))_2,(\varphi_U(q))_3).$$

一般地,记

$$u_j(q)\equiv(\varphi_U(q))_je_j\equiv(e_j\circ\varphi_U)(q),\quad j=1,2,\cdots,m,$$

于是

$$((\varphi_U(q))_1,\cdots,(\varphi_U(q))_m)=\sum_{j=1}^m(\varphi_U(q))_j\,e_j=\sum_{j=1}^m(e_j\circ\varphi_U)(q).\qquad(6.1.5)$$

注意,这里的$\{e_1,\cdots,e_m\}$是$\mathbb{R}^m$ 中取定的一个坐标系,并不一定是通常使用的标准直交系.

由于流形的光滑性,知在(6.1.4)式中,$u_j\in C_p^\infty(M)$,于是

$$u_j\in C_p^\infty(M)\Rightarrow[u_j]\in\mathfrak{F}\Rightarrow[u_j]^*\in\mathfrak{F}/\sim\Rightarrow(\mathrm{d}u_j)\equiv[u_j]^*\in T_p^*,$$

此即$(\mathrm{d}u_j)_p=[u_j]^*\in T_p^*$. 我们要证明

$$\{(\mathrm{d}u_j)_p,\ 1\leqslant j\leqslant m\}\qquad(6.1.6)$$

是 T_p^* 的一个基,即要证明两件事:

① $\forall(\mathrm{d}f)_p\in T_p^*$,都可由$(\mathrm{d}u_j)_p(1\leqslant j\leqslant m)$线性表示;

② $(\mathrm{d}u_j)_p(1\leqslant j\leqslant m)$是线性无关的.

首先,$\forall(\mathrm{d}f)_p\in T_p^*$,则 $f\circ\varphi_U^{-1}$ 是定义在$\mathbb{R}^m$ 中某个开集 $V(\varphi_U(q)\in V)$上的光滑函数,对于$(x_1,x_2,\cdots,x_m)\in V\subset\mathbb{R}^m$,令

$$F(x_1,x_2,\cdots,x_m)\equiv f\circ\varphi_U^{-1}(x_1,x_2,\cdots,x_m),$$

于是 $f(q)=F(u_1,u_2,\cdots,u_m)$. 利用定理 6.1.2,得到

$$(\mathrm{d}f)_p=\sum_{k=1}^m\left(\frac{\partial F}{\partial u_k}\right)_{(u_1(p),\cdots,u_m(p))}(\mathrm{d}u_k)_p,$$

所以$(\mathrm{d}f)_p$ 是$(\mathrm{d}u_j)_p(1\leqslant j\leqslant m)$的线性组合,这就证明了①.

其次,证$(\mathrm{d}u_j)_p(1\leqslant j\leqslant m)$是线性无关的. 若不然,有一组不全为 0 的实数 $\alpha_j,1\leqslant j\leqslant m$,使得

$$\sum_{j=1}^{m}\alpha_j(\mathrm{d}u_j)_p=0\Rightarrow\sum_{j=1}^{m}\alpha_j[u_j]\in\mathfrak{J}_p$$

$$\Rightarrow\quad\forall\gamma\in\Gamma_p\text{ 蕴含 }\left\langle\!\!\left\langle\gamma,\sum_{j=1}^{m}\alpha_j[u_j]\right\rangle\!\!\right\rangle=0\text{ 蕴含 }\sum_{j=1}^{m}\alpha_j\frac{\mathrm{d}(u_j\circ\gamma(t))}{\mathrm{d}t}\bigg|_{t=0}=0\qquad(6.1.7)$$

$\Rightarrow$ 取一组 $\lambda_k\in\Gamma_p,1\leqslant k\leqslant m$,满足 $u_j\circ\lambda_k(t)=u_j(p)+t\delta_{kj}$,其中 $\delta_{kj}=\begin{cases}1, & j=k,\\0, & j\neq k\end{cases}$(这组 $\lambda_k\in\Gamma_p,1\leqslant k\leqslant m$,实际上是取的通过 p 点的直线,它们当然是"曲线"的特例)

$$\Rightarrow\quad\frac{\mathrm{d}(u_j\circ\lambda_k(t))}{\mathrm{d}t}\bigg|_{t=0}=\frac{\mathrm{d}(u_j(p))}{\mathrm{d}t}\bigg|_{t=0}+\frac{\mathrm{d}(t\delta_{kj})}{\mathrm{d}t}=0+\delta_{kj}$$

$\Rightarrow$ 在(6.1.7)式中,令 $\gamma(t)=\lambda_k(t)$,便得到

$$\gamma=\lambda_k\in\Gamma_p\text{ 蕴含 }\sum_{j=1}^{m}\alpha_j\frac{\mathrm{d}(u_j\circ\lambda_k(t))}{\mathrm{d}t}\bigg|_{t=0}=\sum_{j=1}^{m}\alpha_j\delta_{kj}=0,\quad k=1,2,\cdots,m.$$

由此得到 $\alpha_j=0,1\leqslant j\leqslant m$,这与 α_j 不全为 0 矛盾,从而②成立,即$(\mathrm{d}u_j)_p(1\leqslant j\leqslant m)$线性无关. 定理得证.

我们称(6.1.6)式决定的$\{(\mathrm{d}u_j)_p,1\leqslant j\leqslant m\}$为 T_p^* 的(关于光滑流形 M 的局部坐标系$\{(U,u_j)\}$的)**自然基**.

7. 切空间 $T_p=\Gamma_p/\sim$ ——参数曲线空间 Γ_p 中的等价类

由配合的定义(6.1.2)式,对于 $[f]\in\mathfrak{F}_p$ 与 $\gamma\in\Gamma_p$,称

$$\langle\langle\gamma,[f]\rangle\rangle=\frac{\mathrm{d}(f\circ\gamma)}{\mathrm{d}t}\bigg|_{t=0},\quad t\in(-\delta,\delta)$$

为$[f]\in\mathfrak{F}_p$ 与 $\gamma\in\Gamma_p$ 的配合. 由此配合,可定义两种等价关系

① 在 $\mathfrak{F}_p$ 中 —— $[f],[g]\in\mathfrak{F}_p$

$$[g]\sim[f]\quad\Leftrightarrow\quad\langle\langle\gamma,[f]\rangle\rangle=\langle\langle\gamma,[g]\rangle\rangle,\ \forall\gamma\in\Gamma_p;\qquad(6.1.8)$$

从而得到余切空间 $T_p^*=\mathfrak{F}_p/\sim$;

② 在 Γ_p 中 —— $\gamma,\gamma'\in\Gamma_p$

$$\gamma\sim\gamma'\quad\Leftrightarrow\quad\langle\langle\gamma,[f]\rangle\rangle=\langle\langle\gamma',[f]\rangle\rangle,\ \forall[f]\in\mathfrak{F}_p,$$

因$[f]^*=(\mathrm{d}f)_p$ 与$[f]$的选取无关,上式成为

$$\gamma\sim\gamma'\quad\Leftrightarrow\quad\langle\langle\gamma,(\mathrm{d}f)_p\rangle\rangle=\langle\langle\gamma',(\mathrm{d}f)_p\rangle\rangle,\quad\forall(\mathrm{d}f)_p\in T_p^*,\qquad(6.1.9)$$

从而得到切空间 $T_p=\Gamma_p/\sim$.

定义 6.1.9(**Γ_p 中的等价类,切空间**) 设 $T_p^*=\mathfrak{F}_p/\sim$为余切空间,我们定义:

(1) 参数曲线空间 $\Gamma_p=\{\gamma:(-\delta,\delta)\to M\}$中的等价关系

在流形 M 上过点 $p\in M$ 的参数曲线空间 Γ_p 中,定义等价关系$\sim$:

$$\gamma,\gamma'\in\Gamma_p,\gamma\sim\gamma'\Leftrightarrow\langle\langle\gamma,(\mathrm{d}f)_p\rangle\rangle=\langle\langle\gamma',(\mathrm{d}f)_p\rangle\rangle,\ \forall(\mathrm{d}f)_p\in T_p^*.$$

(2) 切空间 $T_p=\Gamma_p/\sim$

在 Γ_p 中的等价关系(6.1.9)化为

$$\gamma,\gamma'\in\Gamma_p,\ \gamma\sim\gamma'\quad\Leftrightarrow\quad\langle\langle\gamma-\gamma',(\mathrm{d}f)_p\rangle\rangle=0,\ \forall(\mathrm{d}f)_p\in T_p^*.$$

记等价类为

$$[\gamma]=\{\gamma'\in\Gamma_p:\gamma'\sim\gamma\},$$

并记等价类的全体为

$$T_p=\Gamma_p/\sim=\{[\gamma]:\gamma\in\Gamma_p\}.$$

在此等价类集 $T_p=\{[\gamma]:\gamma\in\Gamma_p\}$ 中，定义加法与数乘，使得 $T_p=\{[\gamma]:\gamma\in\Gamma_p\}$ 成为线性空间，称 $T_p=\Gamma_p/\sim$ 为流形 M 在点 $p\in M$ 的**切空间(tangent space)**.

定理 6.1.5(切空间的性质) 设 M 为 m 维光滑流形，则

(1) $T_p=\{[\gamma]:\gamma\in\Gamma_p\}$ 是线性空间；

(2) T_p 是余切空间 T_p^* 的对偶空间，亦即，T_p 是 T_p^* 的线性泛函空间 $(T_p^*)^*=T_p$，且 $\dim T_p^*=m=\dim T_p$；

(3) 对于 T_p 中的每个元 $[r]\in T_p$，可以定义 $\langle[\gamma],(\mathrm{d}f)_p\rangle\equiv\langle\langle\gamma,(\mathrm{d}f)_p\rangle\rangle$，$\forall\gamma\in[\gamma]\in T_p$，$\forall(\mathrm{d}f)_p\in T_p^*$，使得 $\langle[\gamma],(\mathrm{d}f)_p\rangle$ 是 $T_p\times T_p^*$ 上的双线性式.

证 (1) $T_p=\{[\gamma]:\gamma\in\Gamma_p\}$ 是线性空间，因为

$$[\gamma_1],[\gamma_2]\in T_p\Rightarrow[\gamma_1]+[\gamma_2]=[\gamma_1+\gamma_2],$$

$$[\gamma]\in T_p,\alpha\in\mathbb{R}\Rightarrow\alpha[\gamma]=[\alpha\gamma],$$

所以 T_p 构成实线性空间.

(2) 为证 $(T_p^*)^*=T_p$，取光滑流形 M 的点 $q\in U$ 的局部坐标，如(6.1.4)式

$$(u_1(q),\cdots,u_m(q))=((\varphi_U(q))_1,\cdots,(\varphi_U(q))_m)=((e_1\circ\varphi_U)(q),\cdots,(e_m\circ\varphi_U)(q)).$$

设 M 上一条光滑参数曲线 $\gamma\in\Gamma_p$ 由函数 $u_j=u_j(t)(1\leqslant j\leqslant m)$ 给出，$p=(u_j(0))_{1\leqslant j\leqslant m}$ 是点 $p\in M$ 的局部坐标. 故由定理 6.1.1 证明中计算 $\langle\langle\gamma,[f]\rangle\rangle$ 的思路，以及定理 6.1.2，可推导如下：

$$\forall[\gamma]\in T_p,\forall(\mathrm{d}f)_p\in T_p^*\ \text{有}\quad\langle[\gamma],(\mathrm{d}f)_p\rangle\equiv\langle\langle\gamma,(\mathrm{d}f)_p\rangle\rangle$$

$$\Rightarrow\ \langle[\gamma],(\mathrm{d}f)_p\rangle=\sum_{j=1}^{m}a_j\xi_j,\text{其中}$$

$$a_j=\left(\frac{\partial(f\circ\varphi_U^{-1})}{\partial u_j}\right)_{\varphi_U(p)},\ \xi_j=\left(\frac{\mathrm{d}u_j}{\mathrm{d}t}\right)_{t=0}=\left(\frac{\mathrm{d}(u_j\circ\gamma(t))}{\mathrm{d}t}\right)_{t=0},\quad 1\leqslant j\leqslant m\quad(6.1.10)$$

$\Rightarrow$ 上式表明系数 $a_j=\left(\dfrac{\partial(f\circ\varphi_U^{-1})}{\partial u_j}\right)_{\varphi_U(p)}$ 恰是余切矢量 $(\mathrm{d}f)_p\in T_p^*$ 关于自然基 $\{\xi_j:1\leqslant j\leqslant m\}\equiv\{(\mathrm{d}u_j)_p,1\leqslant j\leqslant m\}$ 的分量

$\Rightarrow\ \langle[\gamma],(\mathrm{d}f)_p\rangle=\sum_{j=1}^{m}a_j\xi_j$ 是 T_p^* 到数域$\mathbb{R}$的线性映射,亦即,T_p^* 到$\mathbb{R}$的映射 $(\mathrm{d}f)_p\in T_p^*\to\langle[\gamma],(\mathrm{d}f)_p\rangle\in\mathbb{R}$ 是一个线性泛函,当 $(\mathrm{d}f)_p\in T_p^*$ 给定时,它由分量

$$\xi_j=\left(\frac{\mathrm{d}u_j}{\mathrm{d}t}\right)_{t=0}=\left(\frac{\mathrm{d}(u_j\circ\gamma(t))}{\mathrm{d}t}\right)_{t=0},\quad 1\leqslant j\leqslant m$$

完全确定;而 $a_j=\left(\frac{\partial(f\circ\varphi_U^{-1})}{\partial u_j}\right)_{\varphi_U(p)}(1\leqslant j\leqslant m)$ 则由 $(\mathrm{d}f)_p\in T_p^*$ 确定

$\Rightarrow\ \forall\gamma\in[\gamma]\in T_p$,可确定一组 $\{\xi_j:1\leqslant j\leqslant m\}$,也就确定一个线性泛函. 将其记为

$$F_{[\gamma]}\leftrightarrow\xi_j=\left(\frac{\mathrm{d}(u_j\circ\gamma(t))}{\mathrm{d}t}\right)_{t=0},\quad 1\leqslant j\leqslant m.$$

$\Rightarrow\ \left\{F_{[\gamma]}(\cdot)=\sum_{j=1}^{m}a_j\xi_j=\langle\langle[\gamma],\cdot\rangle\rangle:\forall[\gamma]\in T_p\right\}\leftrightarrow T_p$ 是 T_p^* 上的线性泛函的全体,故 $(T_p^*)^*=T_p$ 得证.

(3) 显然. 定理得证.

定义 6.1.10(切矢量) 设 $T_p=\{[\gamma]:\gamma\in\Gamma_p\}$ 为流形 $(M,\mathbb{A})$ 在点 $p\in M$ 的切空间. 切空间 T_p 中的元 $[\gamma]$ 称为过点 p 的**切矢量(tangent vector)**.

切矢量的几何意义 对于两个切矢量 $[\gamma],[\gamma']\in T_p$,若 $\gamma,\gamma'\in\Gamma_p$ 分别有局部坐标

$$\gamma\in\Gamma_p:(-\delta,\delta)\to M:\quad (u_j(t))=(u_1(t),\cdots,u_m(t)),\quad t\in(-\delta,\delta),$$
$$\gamma'\in\Gamma_p:(-\delta,\delta)\to M:\quad (u_j'(t))=(u_1'(t),\cdots,u_m'(t)),\quad t\in(-\delta,\delta),$$

则 $[\gamma]\sim[\gamma']\Leftrightarrow\left(\frac{\mathrm{d}u_j}{\mathrm{d}t}\right)_{t=0}=\left(\frac{\mathrm{d}u_j'}{\mathrm{d}t}\right)_{t=0}$, $1\leqslant j\leqslant m$. 故 $[\gamma]$ 与 $[\gamma']$ 等价是指两条光滑参数曲线 $\gamma,\gamma'\in\Gamma_p$ 在点 $p\in M$ 有相同的切矢量,如图 6.1.10 所示. 因此,流形 M 在点 p 的切矢量恰好是在点 p 的有相同切矢量的参数曲线的全体所成的集合.

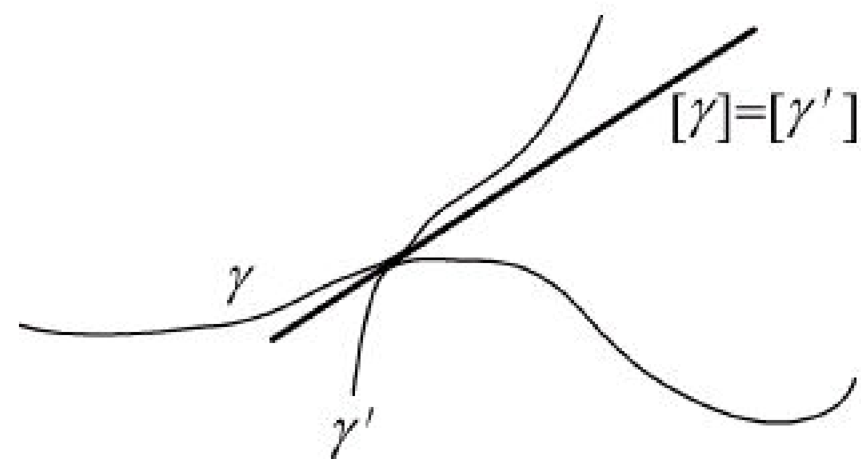


图 6.1.10 切矢量

根据以上讨论,记 $X=X_p=[\gamma]$,则定理 6.1.5(3) 给出

$$\langle[\gamma],(\mathrm{d}f)_p\rangle\equiv\langle X,(\mathrm{d}f)_p\rangle,\quad X=[\gamma]\in T_p,\ (\mathrm{d}f)_p\in T_p^*$$

是双线性式.

下面的定义给出此双线性式的意义.

定义 6.1.11(光滑函数的微分) 光滑流形 $(M,\mathbb{A})$ 在点 $p\in M$ 的切空间 T_p 中的元 $[\gamma]$,即过点 p 的切矢量,记为 $X=X_p=[\gamma]$, $X\in T_p$;余切空间 T_p^* 中的元 $[f]^*$,即过点 p 的余切矢量,记为 $(\mathrm{d}f)_p=[f]^*$. 把 $(\mathrm{d}f)_p\in T_p^*$ 称为 $f\in C_p^\infty(M)$ 在点 p 的**微分(differential)**,并且把双线性式

$$Xf\equiv\langle X,(\mathrm{d}f)_p\rangle,\quad X\in T_p,\ (\mathrm{d}f)_p\in T_p^*$$

称为函数 f 沿切矢量 X 的**方向导数(directional derivative)**.

若微分$(\mathrm{d}f)_p=0$,则称点 p 是函数 f 的**临界点**(**critical point**).

定理 6.1.6 方向导数有如下性质: $\forall X\in T_p, f,g\in C_p^\infty(M),\alpha,\beta\in\mathbb{R}$,则

(1) $X(\alpha f+\beta g)=\alpha Xf+\beta Xg$;

(2) $X(f\cdot g)=f(p)Xg+g(p)Xf$.

于是,$X=X_p:C_p^\infty(M)\to\mathbb{R}$ 是 $C_p^\infty(M)$上的实函数.

证 由定理 6.1.2 与定理 6.1.3,就可得到此定理的结论.

注 将切矢量 $X=X_p\in T_p$ 看作算子(方向导数算子),定理 6.1.6 说明,$X\in T_p$ 是 $C_p^\infty(M)$ 上的线性算子.

8. T_p^* 与 T_p 的自然基

考虑 $\forall X=X_p=[\gamma]\in T_p$, $\forall(\mathrm{d}f)_p\in T_p^*$,式 $Xf\equiv\langle X,(\mathrm{d}f)_p\rangle$是双线性式. 于是,由光滑流形$(M,\mathbb{A})$的局部坐标系$\{u_j(q):1\leqslant j\leqslant m\}$,$q\in U_p\subset M$,

$$(u_1(q),\cdots,u_m(q))=((\varphi_U(q))_1,\cdots,(\varphi_U(q))_m)=((e_1\circ\varphi_U)(q),\cdots,(e_m\circ\varphi_U)(q)),$$

得到 T_p^* 的自然基$\{(\mathrm{d}u_j)_p:1\leqslant j\leqslant m\}$,然后取光滑参数曲线 $\lambda_k\in\Gamma_p$,$1\leqslant k\leqslant m$,满足

$$\langle[\lambda_k],(\mathrm{d}u_j)_p\rangle=\delta_{kj}=\begin{cases}1, & j=k,\\ 0, & j\neq k,\end{cases}\tag{6.1.11}$$

参见定理 6.1.4 证明的第二步,取 $u_j\circ\lambda_k(t)=u_j(p)+t\delta_{kj}$,其中 δ_{kj} 如(6.1.11)式定义,所以$\{[\lambda_k],1\leqslant k\leqslant m\}\subset T_p$ 是自然基$\{(\mathrm{d}u_j)_p:1\leqslant j\leqslant m\}\subset T_p^*$ 的对偶基.

再来进一步理解满足(6.1.11)式的切矢量$[\lambda_k]\in T_p$,$1\leqslant k\leqslant m$. 由于定理 6.1.5(3),将式

$$\langle[\gamma],(\mathrm{d}f)_p\rangle\equiv\langle\langle\gamma,(\mathrm{d}f)_p\rangle\rangle,\quad\forall\gamma\in[\gamma]\in T_p,\ \forall(\mathrm{d}f)_p\in T_p^*$$

取为 $\langle[\gamma],(\mathrm{d}f)_p\rangle=\sum\limits_{j=1}^m a_j\xi_j$,其中

$$a_j=\left(\frac{\partial(f\circ\varphi_U^{-1})}{\partial u_j}\right)_{\varphi_U(p)},\quad \xi_j=\left(\frac{\mathrm{d}u_j}{\mathrm{d}t}\right)_{t=0}=\left(\frac{\mathrm{d}(u_j\circ\gamma(t))}{\mathrm{d}t}\right)_{t=0},\quad 1\leqslant j\leqslant m;$$

另一方面,$(\mathrm{d}f)_p=\sum\limits_{j=1}^m\left(\frac{\partial f}{\partial u_j}\right)_p\cdot(\mathrm{d}u_j)_p$,其中$\left(\frac{\partial f}{\partial u_j}\right)_p=\left(\frac{\partial(f\circ\varphi_U^{-1})}{\partial u_j}\right)_p=a_j$.

于是,利用(6.1.11)式,有

$$\langle[\lambda_k],(\mathrm{d}f)_p\rangle=\left\langle[\lambda_k],\sum_{j=1}^m\left(\frac{\partial f}{\partial u_j}\right)_p\cdot(\mathrm{d}u_j)_p\right\rangle=\left(\frac{\partial f}{\partial u_k}\right)_p=\left(\frac{\partial(f\circ\varphi_U^{-1})}{\partial u_k}\right)_p,$$

所以,切矢量$[\lambda_k]\in T_p$,$1\leqslant k\leqslant m$,对于由函数芽$[f]$产生的$[f]^*=(\mathrm{d}f)_p$ 的"作用",实际上是偏微分算子$[\lambda_k]=\left(\frac{\partial}{\partial u_k}\right)_p$ 的作用: $\langle[\lambda_k],(\mathrm{d}f)_p\rangle=\left(\frac{\partial f}{\partial u_k}\right)_p=\left(\frac{\partial(f\circ\varphi_U^{-1})}{\partial u_k}\right)_p$,于是,(6.1.11)式

可写成

$$\left\langle \frac{\partial}{\partial u_k}\bigg|_p, (\mathrm{d}u_j)_p \right\rangle = \delta_{kj} = \begin{cases} 1, & j = k, \\ 0, & j \neq k. \end{cases} \tag{6.1.12}$$

对光滑流形$(M,\mathbb{A})$,定义余切空间 T_p^* 与切空间 T_p 之后,设 T_p^* 的自然基为

$$\{(\mathrm{d}u_j)_p, 1 \leqslant j \leqslant m\} \subset T_p^*,$$

则其对偶基(即 T_p 的自然基)为

$$\left\{\frac{\partial}{\partial u_j}\bigg|_p, 1 \leqslant j \leqslant m\right\} \subset T_p.$$

9. T_p 与 T_p^* 中的元的表示

通常,用自然基表示

$$\forall X \in T_p \Rightarrow X = \sum_{j=1}^m \xi_j \frac{\partial}{\partial u_j}, \quad \text{其中 } \xi_j = \frac{\mathrm{d}(u_j \circ \gamma)}{\mathrm{d}t}, X = [\gamma];$$

$$\forall a \in T_p^* \Rightarrow a = (\mathrm{d}f)_p = \sum_{k=1}^m a_k \mathrm{d}u_k, \quad \text{其中 } a_k = \frac{\partial f}{\partial u_k}.$$

当 T_p^* 有另一组局部坐标系

$$\{(\mathrm{d}\,\tilde{u}_j)_p, 1 \leqslant j \leqslant m\} \subset T_p^*$$

时,X 与 a 分别表示为

$$\forall X \in T_p \Rightarrow X = \sum_{j=1}^m \tilde{\xi}_j \frac{\partial}{\partial \tilde{u}_j}, \quad \text{其中} \tilde{\xi}_j = \frac{\mathrm{d}(\tilde{u}_j \circ \gamma)}{\mathrm{d}t}, X = [\gamma];$$

$$\forall a \in T_p^* \Rightarrow a = (\mathrm{d}f)_p = \sum_{k=1}^m \tilde{a}_k \mathrm{d}\tilde{u}_k, \quad \text{其中} \tilde{a}_k = \frac{\partial f}{\partial \tilde{u}_k}.$$

不难推出如下变换关系:

$$\tilde{\xi}_j = \sum_{k=1}^m \xi_k \frac{\partial \tilde{u}_j}{\partial u_k}, \quad a_k = \sum_{j=1}^m \tilde{a}_j \frac{\partial \tilde{u}_j}{\partial u_k}, \tag{6.1.13}$$

其中

$$\frac{\partial \tilde{u}_j}{\partial u_k} = \frac{\partial(\varphi_{\tilde{U}} \circ \varphi_U^{-1})_j}{\partial u_k} \tag{6.1.14}$$

正好给出坐标变换 $\varphi_{\tilde{U}} \circ \varphi_U^{-1}$ 的 Jacobi 矩阵中的元.

再次强调,我们用$\{(U,u_j)\}$表示光滑流形$(M,\mathbb{A})$的一个局部坐标系.

10. 光滑流形之间的余切空间与切空间

前几节中,研究光滑流形$(M,\mathbb{A})$上一点 $p\in M$ 的余切空间、切空间,现在转向整个流形$(M,\mathbb{A})$,$\forall\, p\in M$ 的余切空间与切空间的全体.

1）光滑流形之间映射 $F:M\to N$ 的微分

定义 6.1.12（**光滑流形之间的映射 F 的微分**） 设 $(M,\mathbb{A})$ 与 $(N,\mathbb{B})$ 分别为 m 维与 n 维光滑流形，$F:M\to N$ 为光滑映射，$p\in M$，像点 $q=F(p)\in N$. 记余切空间为 $T_p^*(M)$ 与 $T_q^*(N)$. 定义映射 $F^*:T_q^*(N)\to T_p^*(M)$，$p\in M$，$q\in N$，满足

$$F^*((\mathrm{d}f)_q)=\mathrm{d}(f\circ F),\quad (\mathrm{d}f)_q\in T_q^*(N),$$

称 $F^*:T_q^*(N)\to T_p^*(M)$ 为**映射 $F:M\to N$ 的微分**.

注 $\forall(\mathrm{d}f)_q\in T_q^*(N)\Rightarrow F^*((\mathrm{d}f)_q)=\mathrm{d}(f\circ F)\in T_p^*(M)$，亦即，$F^*$ 把 $T_q^*(N)$ 中的微分 $(\mathrm{d}f)_q$ 映到 $T_p^*(M)$ 中的微分 $\mathrm{d}(f\circ F)$，由 $F:M\to N$，$(\mathrm{d}f)_q\in T_q^*(N)$ 知道，$f\in C_q^\infty(N)$，故 $f\circ F:M\to\mathbb{R}$ 为 M 上的 C^∞ 函数，$\mathrm{d}(f\circ F)$ 当然有意义，因此定义映射

$$F^*:T_q^*(N)\to T_p^*(M)$$

为 $F:M\to N$ 的微分是有意义的. 微分 F^* 也称为映射 $F:M\to N$ 的余切映射.

2）光滑流形之间映射 $F:M\to N$ 的切映射

定义 6.1.13（**光滑流形之间的映射 F 的切映射**） 设 $(M,\mathbb{A})$ 与 $(N,\mathbb{B})$ 分别为 m 维与 n 维光滑流形，$F:M\to N$ 为光滑映射，$p\in M$，像点 $q=F(p)\in N$. 记余切空间为 $T_p^*(M)$ 与 $T_q^*(N)$. 设 $F^*:T_q^*(N)\to T_p^*(M)$ 是定义 6.1.12 中的 $F:M\to N$ 的微分. 记映射

$$F_*:T_p(M)=(T_p^*(M))^*\to T_q(N)=(T_q^*(N))^*$$

为 $F^*:T_q^*(N)\to T_p^*(M)$ 的共轭映射，亦即

$$\langle F_*X,a\rangle=\langle X,F^*a\rangle,\quad X\in T_p(M),\ a\in T_q^*(N),$$

微分 $F^*:T_q^*(N)\to T_p^*(M)$ 的共轭映射 $F_*:T_p(M)\to T_q(N)$ 称为由**映射 $F:M\to N$ 诱导的切映射**，简称为**映射 $F:M\to N$ 的切映射**.

注 此概念的理解如下.

(1) 映射 $F:M\to N\Rightarrow$ 微分 $F^*:T_q^*(N)\to T_p^*(M)$ 满足

$$F^*((\mathrm{d}f)_q)=(\mathrm{d}(f\circ F))_p,\quad (\mathrm{d}f)_q\in T_q^*(N).$$

由于 $(T_p^*(M))^*=T_p(M)$，$(T_q^*(N))^*=T_q(N)$，于是，对于 $F^*:T_q^*(N)\to T_p^*(M)$，$F^*((\mathrm{d}f)_q)=(\mathrm{d}(f\circ F))_p$，$\forall(\mathrm{d}f)_q\in T_q^*(N)$，作用的关系为

$F:M\to N \longrightarrow f\circ F:M\to\mathbb{R} \longrightarrow (\mathrm{d}(f\circ F))_p\in T_p^*(M)$

$\uparrow$

$f:N\to\mathbb{R}\leftarrow f\in C_q^\infty(N)$

$\downarrow$

有意义

$(\mathrm{d}f)_q\in T_q^*(N)$

$\downarrow$

$F^*:T_q^*(N)\to T_p^*(M)\longrightarrow F^*((\mathrm{d}f)_q)\in T_p^*(M)$

(2) F 的微分 $F^*:T_q^*(N)\to T_p^*(M)$ ⇒ F 的切映射 $F_*:T_p(M)\to T_q(N)$，满足

$$\langle F_*X,a\rangle=\langle X,F^*a\rangle,\quad X\in T_p(M),\ a\in T_q^*(N).$$

根据(1)，$F^*:T_q^*(N)\to T_p^*(M)$，故 $F^*:\forall a\in T_q^*(N)\to F^*a\in T_p^*(M)$，因而

$$\forall X\in T_p(M),\forall F^*a\in T_p^*(M)\Rightarrow\langle X,F^*a\rangle\in\mathbb{R}$$

有意义；于是，可以定义左边的算子 $F_*:T_p(M)\to T_q(N)$：$\forall X\in T_p(M)$，每个 F_*X 视为 $T_q(N)$中的元.

对于$(T_q(N))^*=T_q^*(N)$中的元 $a\in T_q^*(N)$，$\langle F_*X,a\rangle$有定义，而且属于$\mathbb{R}$，于是，让$\langle F_*X,a\rangle$与$\langle X,F^*a\rangle$相等，就定义了切映射 $F_*:T_p(M)\to T_q(N)$.

为什么称 F_* 为 $F:M\to N$ 的切映射呢？请读者自己与高等数学中的函数相切的概念相比较.

图 6.1.11 给出了切空间与切映射的一个简单的图解.

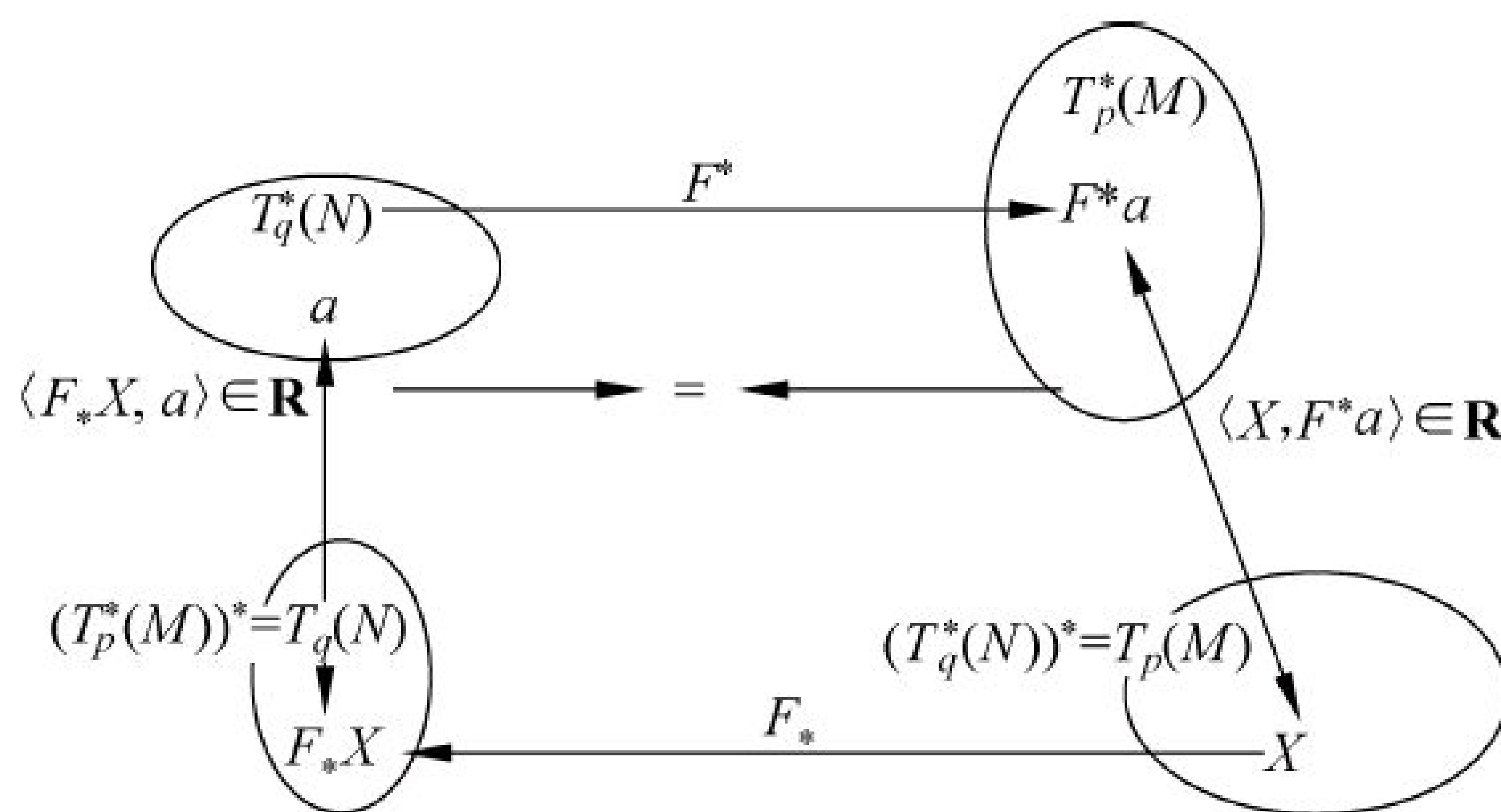


图 6.1.11 切映射、切空间

最后来求 $F^*:T_q^*(N)\to T_p^*(M)$与 $F_*:T_p(M)\to T_q(N)$在自然基之下的变换矩阵.

设 $p\in M$ 的开邻域中的局部坐标为$(u_j)_{1\leqslant j\leqslant m}$，像点 $f(p)=q\in N$ 的开邻域中的局部坐标为$(v_k)_{1\leqslant k\leqslant n}$，假设由此构成 $T_p^*(M)$中的自然基$\{(\mathrm{d}u_j):1\leqslant j\leqslant m\}$与 $T_p^*(N)$中的自然基$\{(\mathrm{d}v_k):1\leqslant k\leqslant n\}$，于是，$F:M\to N$ 在 $p\in M$ 点的开邻域中可表示为

$$v_k=F_k(u_1,\cdots,u_m),\quad 1\leqslant k\leqslant n,$$

而 $F^*:T_q^*(N)\to T_p^*(M)$在 $T_p^*(N)$中的自然基$\{(\mathrm{d}v_k):1\leqslant k\leqslant n\}$上的作用可记为

$$F^*((\mathrm{d}v_k)_q)=\mathrm{d}((v_k\circ F)_p)=\sum_{j=1}^{m}\left(\frac{\partial F_k}{\partial u_j}\right)_p\mathrm{d}u_j,\quad k=1,\cdots,n.$$

此式表明，F^* 将 $T_p^*(N)$中的自然基$\{(\mathrm{d}v_k):1\leqslant k\leqslant n\}$变换到 $T_p^*(M)$中的自然基$\{(\mathrm{d}u_j):1\leqslant j\leqslant m\}$之间的变换 Jacobi 矩阵是$\left[\left(\frac{\partial F_k}{\partial u_j}\right)_p\right]_{n\times m}$.

同理，相应的切映射 $F_*:T_p(M)\to T_q(N)$在 $T_p(M)$中的自然基$\left\{\frac{\partial}{\partial u_j}\right\}_{1\leqslant j\leqslant m}$上的作用是

$$\left\langle F_*\left(\frac{\partial}{\partial u_j}\right),\mathrm{d}v_k\right\rangle=\left\langle\frac{\partial}{\partial u_j},F^*(\mathrm{d}v_k)\right\rangle=\sum_{s=1}^{m}\left\langle\frac{\partial}{\partial u_j},\mathrm{d}u_s\right\rangle\left(\frac{\partial F_k}{\partial u_s}\right)_p$$

$$= \left\langle \sum_{t=1}^{n} \left(\frac{\partial F_t}{\partial u_j}\right)_p \frac{\partial}{\partial v_t}, \mathrm{d}v_k \right\rangle,$$

此即

$$F_* \left(\frac{\partial}{\partial u_j}\right) = \sum_{t=1}^{n} \left(\frac{\partial F_t}{\partial u_j}\right)_p \frac{\partial}{\partial v_t}, \quad j = 1,2,\cdots,m.$$

因此,切映射 $F_*: T_p(M) \to T_q(N)$ 将 $T_p(M)$ 中的自然基 $\left\{\frac{\partial}{\partial u_j}\right\}_{1\leqslant j\leqslant m}$ 变换到 $T_q(N)$ 中的自然基 $\left\{\frac{\partial}{\partial v_k}\right\}_{1\leqslant k\leqslant n}$ 的变换 Jacobi 矩阵是 $\left[\left(\frac{\partial F_k}{\partial u_j}\right)_p\right]_{m\times n}$.

11. 切矢量场

6.1.1 节与 6.1.2 节中,我们在 T_2 型拓扑空间 (M,τ) 上建立了 C^∞ 微分流形 $(M,\tau,\mathbb{A})$,流形上的余切空间 $T_p^*(M)$ 与切空间 $T_p(M)$.

定理 6.1.6 告诉我们,切矢量 $X_p \in T_p$ 是定义在 $C_p^\infty(M)$ 上的实函数,即

$$X_p: C_p^\infty(M) \to \mathbb{R}.$$

如果对于流形 M 上的每一点 $p\in M$,指定 M 在点 p 的一个切矢量 X_p,则称

$$X = \{X_p : p \in M\}$$

是**流形 M 上的切矢量场**.

若 $f\in C^\infty(M)$,令 $(Xf)(p)=X_p f$,则 Xf 是流形 M 上的实函数.

定义 6.1.14(光滑切矢量场) 设 $X=\{X_p:p\in M\}$ 是光滑流形 M 上的一个切矢量场,若对于任意的 $f\in C^\infty(M)$,仍有 $Xf\in C^\infty(M)$,则称 X 是流形 M 上的**光滑切矢量场**. 若切矢量场 $X=\{X_p:p\in M\}$ 中的每个切矢量都是光滑的,则其为光滑切矢量场. 于是,光滑切矢量场 $X=\{X_p:p\in M\}$ 是 $C^\infty(M)$ 到 $C^\infty(M)$ 的一个算子. 其性质如下:

(1) $X(\alpha f+\beta g)=\alpha(Xf)+\beta(Xg)$;

(2) $X(f\cdot g)=f\cdot(Xg)+g\cdot(Xf)$,

其中 $f,g\in C^\infty(M)$,$\alpha,\beta\in\mathbb{R}$.

引理 6.1.1 设 $X=\{X_p:p\in M\}$ 是光滑流形 M 上的光滑切矢量场,则对 M 的任一个非空开子集 U,X 在 U 上的限制 $X|_U$ 是开子流形 U 上的光滑切矢量场.

证 只需证明,$\forall f\in C^\infty(U)$,$X|_U f$ 仍然是 U 上的光滑函数.

(1) 任取一点 $p\in U$,则存在点 p 的坐标域 V,$p\in V$,使得 $\bar{V}$ 是紧致的,且 $\bar{V}\subset U$. 于是,在流形 M 上存在光滑函数 $h: M\to\mathbb{R}$,使得

① $0\leqslant h\leqslant 1$;　　② $h(q)=\begin{cases}1, & p\in V,\\ 0, & p\notin V.\end{cases}$

其证明方法与局部紧拓扑空间的 Urysohn 引理证法类似.

(2) 令 $F(x)=\begin{cases} f(x)\cdot h(x), & x\in U,\\ 0, & x\notin U,\end{cases}$ 则由 f 与 h 的光滑性,得到 $F\in C^\infty(M)$,$F|_V=f|_V$. 由假设,X 是光滑流形 M 上的光滑切矢量场,故 $XF\in C^\infty(M)$. 进而,由

$$(X|_U f)(x)=X|_U f(x)=(Xf)(x),\quad \forall x\in V,$$

因此,函数 $X|_U f$ 在 $q\in U$ 光滑,亦即,$X|_U f$ 是 U 上的光滑函数.

光滑切矢量场有如下坐标表示:

定理 6.1.7 光滑流形 M 上的切矢量场 $X=\{X_p:p\in M\}$ 是光滑切矢量场,当且仅当对于任一点 $p\in M$,存在点 p 的局部坐标系 $\{(U,u_j)\}$,使得 X 限制在 U 上,可表示成

$$X|_U=\sum_{j=1}^m \xi_j\frac{\partial}{\partial u_j},$$

其中 $\xi_j(1\leqslant j\leqslant m)$ 是 U 上的光滑函数.

定义 6.1.15(Poisson 括号积) 设 X,Y 是光滑流形 M 上的两个光滑切矢量场,它们的 Poisson 括号积(简称 Poisson 括号)定义为

$$[X,Y]=XY-YX,$$

亦即,$[X,Y]$ 是作用在 $C^\infty(M)$ 上的算子,满足

$$[X,Y]f=X(Yf)-Y(Xf).$$

定理 6.1.8(Poisson 括号的性质) 设 X,Y,Z 是光滑流形 M 上的光滑切矢量场,则对于 $f,g\in C^\infty(M)$,

(1) $[X,Y](\alpha f+\beta g)=\alpha[X,Y]f+\beta[X,Y]g$, $\alpha,\beta\in\mathbb{R}$;

(2) $[X,Y](f\cdot g)=f\cdot[X,Y]g+g\cdot[X,Y]f$;

(3) $[X,Y]=-[Y,X]$;

(4) $[X+Y,Z]=[X,Z]+[Y,Z]$;

(5) $[fX,gY]=f\cdot(Xg)Y-g\cdot(Yf)X+f\cdot g[X,Y]$;

(6) $[X,[Y,Z]]+[Y,[Z,X]]+[Z,[X,Y]]=0$.

证 只证(5),其余留作习题. 取 $h\in C^\infty(M)$,则由定义,有

$$\begin{aligned}[fX,gY]h&=(fX)((gY)h)-(gY)((fX)h)\\ &=f\cdot X(g\cdot Yh)-g\cdot Y(f\cdot Xh)\\ &=f\cdot(Xg)(Yh)+f\cdot g\cdot X(Yh)-g\cdot(Yf)(Xh)-g\cdot f\cdot Y(Xh)\\ &=(f(Xg)\cdot Y-g(Yf)\cdot X+f\cdot g[X,Y])h.\end{aligned}$$

定理 6.1.9(Poisson 括号的局部坐标表示) 对于光滑流形$(M,\mathbb{A})$,设$\{(U,u_j)\}$为$(M,\mathbb{A})$的一个局部坐标系,并设

$$X|_U=\sum_{j=1}^m \xi_j\frac{\partial}{\partial u_j},\quad Y|_U=\sum_{j=1}^m \eta_j\frac{\partial}{\partial u_j}$$

为两个切矢量场 X 与 Y 的局部坐标表示. 则

$$[X,Y]|_U=[X|_U,Y|_U]=\sum_{j=1}^{m}\sum_{k=1}^{m}\left(\frac{\partial\eta_j}{\partial u_k}-\frac{\partial\xi_j}{\partial u_k}\right)\frac{\partial}{\partial u_j}.$$

证 直接计算 $X=\dfrac{\partial}{\partial u_j}$,得 $\left[\dfrac{\partial}{\partial u_j},\dfrac{\partial}{\partial u_k}\right]=\delta_{kj}=\begin{cases}1, & j=k,\\ 0, & j\neq k,\end{cases}$ 请读者自行验证.

定义 6.1.16(奇点) 设$(M,\mathbb{A})$是光滑流形,X 是$(M,\mathbb{A})$上的光滑切矢量场,若在点 $p\in M$ 有 $X_p=0$,则称点 p 是切矢量场 X 的一个**奇点(singular point)**.

光滑流形 M 的光滑切矢量场 X 的奇点性质是很复杂的,但光滑切矢量场 X 的非奇点的性质比较简单,也很重要. 下面叙述但不加以证明地给出非奇点的局部坐标系的简化定理.

定理 6.1.10 设 M 是光滑流形,X 是$(M,\mathbb{A})$上的光滑切矢量场,若在点 $p\in M$ 有 $X_p\neq0$,则存在点 p 的一个局部坐标系$\{(W,w_j)\}$,使得$X\big|_W=\dfrac{\partial}{\partial w_j}$.

12. 建立光滑流形之间的映射的微分(余切矢量)与切映射流程图

光滑流形 M 与 N 之间的光滑映射 $F:M\to N$

↓

$F:M\to N$ 的微分 —— $F^*:T_q^*(N)\to T_p^*(M)$,定义为

$$F^*((\mathrm{d}f))=\mathrm{d}(f\circ F),\ (\mathrm{d}f)\in T_q^*(N)$$

↓

$F:M\to N$ 的切映射 —— $F_*:T_p(M)\to T_q(N)$,(由 $F^*:T_q^*(N)\to T_p^*(M)$诱导)定义为

$$\langle F_*X,a\rangle=\langle X,F^*a\rangle,\quad X=X_p\in T_p(M),\ a\in T_q^*(N)$$

这些结果在研究流形上的微积分时要用到.

6.1.3 子流形

1. $\mathbb{R}^n$ 中的反函数存在定理

在经典微积分中大家都很熟悉的一个结果:若函数在一点的导数大于零,则函数在该点的一个邻域内是递增的,亦即,一点的微分的性质能决定函数在该点的一个邻域内的性质. 在流形的情形,是否也有类似的性质呢? 答案是肯定的.

首先来回顾一个经典定理.

定理 6.1.11($\mathbb{R}^n$ 中的反函数存在性定理) 设 $W\subset\mathbb{R}^n$ 为开集,$f:W\to\mathbb{R}^n$ 是 W 到$\mathbb{R}^n$ 的光滑映射. 若对 $x_0\in W$,有 $\det\left[\left(\dfrac{\partial f_j}{\partial x_k}\right)_{x_0}\right]\neq0$,则存在 x_0 的邻域 $U\subset W$,使得 $V=f(U)$是

$f(x_0)$在$\mathbb{R}^n$中的邻域,且f在V上有光滑的反函数$g=f^{-1}:V\to U$.

以"流形"的语言,上述定理可理解为映射$f:W\to\mathbb{R}^n$的Jacobi矩阵$\left[\left(\frac{\partial f_j}{\partial x_k}\right)\right]$恰是$f$的切映射$f_*:T_p(\mathbb{R}^n)\to T_q(\mathbb{R}^n)$在自然基$\left\{\frac{\partial}{\partial u_k}\right\}_{1\leqslant k\leqslant n}=\left\{\frac{\partial}{\partial x_k}\right\}_{1\leqslant k\leqslant n}$之下的矩阵,因此行列式$\det\left[\left(\frac{\partial f_j}{\partial x_k}\right)_{x_0}\right]\neq 0$说明线性映射满足如下关系:

$$\begin{array}{ccc} f_*:T_p(W\subset\mathbb{R}^n)\to T_q(\mathbb{R}^n) & \Leftrightarrow & f_*:\mathbb{R}^n\to\mathbb{R}^n \\ \updownarrow \qquad\qquad \updownarrow & & \\ T_p(W\subset\mathbb{R}^n)\approx\mathbb{R}^n \quad T_q(\mathbb{R}^n)\approx\mathbb{R}^n & & \end{array}$$

亦即,f_*是$\mathbb{R}^n$到$\mathbb{R}^n$的同构.

继而,$g=f^{-1}:V\to U$是$f:W\to f(W)\subseteq\mathbb{R}^n$的反函数,是指$g\circ f=I:U\to U$与$f\circ g=I:V\to V$,而且$f$与$g$都是光滑映射. 所以,$f$限制在$U$上,给出了从$U$到$V$上的可微同胚.

反函数存在定理的意义 如果$f:W\to\mathbb{R}^n$的切映射$f_*:T_p(\mathbb{R}^n)\to T_q(\mathbb{R}^n)$在一点$p\in W$是一个同构映射,则$f$在点$p$的一个邻域$U\subset W$中是可微同胚.

重要的是,切映射$F_*:T_p(M)\to T_q(N)$在点$p\in M$的性质,可以决定映射$F:M\to N$在点$p\in M$的一个邻域内的性质.

2. 流形之间切映射的性质

我们有定理6.1.11的推广.

定理 6.1.12 设$(M,\mathbb{A})$与$(N,\mathbb{B})$是两个n维光滑流形,$f:M\to N$是光滑映射. 如果在一点$p\in M$,切映射$f_*:T_p(M)\to T_{f(p)}(N)$是拓扑同构映射,则存在点$p$在$M$中的邻域$U$,使得$f(U)=W\subset N$是点$f(p)\in N$的一个邻域,并且$f|_U:U\to W$是可微同胚.

证 由$f:M\to N$是光滑映射,故可分别取点p在M中的局部坐标卡(U_0,φ)与点$q=f(p)$在N中的局部坐标卡(V_0,ψ),使得$f(U_0)\subset V_0$,并且

$$\tilde{f}=\psi\circ f\circ\varphi^{-1}:\varphi(U_0)\to\psi(V_0)\subset\mathbb{R}^n$$

是光滑映射,显然$\tilde{f}$在点$\varphi(p)$的Jacobi行列式不为零,于是,由定理6.1.11,分别存在点$\varphi(p)$与点$\psi(q)$在$\mathbb{R}^n$中的邻域$\tilde{U}\subset\varphi(U_0)$与$\tilde{V}\subset\varphi(V_0)$,使得$f|_{\tilde{U}}:\tilde{U}\to\tilde{V}$是可微同胚.

令$U=\varphi^{-1}(\tilde{U}),V=\psi^{-1}(\tilde{V})$,则$U,V$分别是点$p,q$在$M,N$中的邻域,并且,映射$f=\psi^{-1}\circ\tilde{f}\circ\varphi:U\to V$是可微同胚. 定理得证.

注 上述定理中的两个流形$(M,\mathbb{A})$,$(N,\mathbb{B})$具有相同的维数,则"切映射f_*在一点同构"等价于"f_*在该点是单一的". 当$(M,\mathbb{A})$是m维光滑流形、$(N,\mathbb{B})$是n维光滑流形时,定理显然成立. 若$f:M\to N$是光滑映射,并且切映射f_*在点$p\in M$是单一映射时,称**切映

射 f_* 在点 p 是非退化的.

当切映射 f_* 是非退化时,必定有 $m\leqslant n$,且 f 的 Jacobi 矩阵在点 $p\in M$ 的秩等于 m.(作为习题)

定理 6.1.13 设$(M,\mathbb{A})$为 m 维光滑流形,$(N,\mathbb{B})$为 n 维光滑流形,$m<n$,且 $f:M\to N$ 是光滑映射.如果切映射 f_* 在点 $p\in M$ 是非退化的,则存在点 p 的局部坐标系$\{(U,u_j)\}$与点 $q=f(p)\in N$ 的局部坐标系$\{(V,v_k)\}$,使得 $f(U)\subset V$,并且映射$f|_U$ 可用局部坐标表示为:对于任意 $x\in U$,有

$$\begin{cases} v_j(f(x)) = u_j(x), & 1\leqslant j\leqslant m, \\ v_s(f(x)) = 0, & m+1\leqslant s\leqslant n. \end{cases} \tag{6.1.15}$$

证明从略.

注 这个定理的意义在于,当切映射 f_* 在点 $p\in M$ 非退化时,必能取得点 p 的局部坐标系$\{(U,u_j)\}$,且有 $u_j(p)=0$,使得映射 $f:M\to N$ 下的像$q=f(p)$的局部坐标系的后 $n-m$ 个坐标为零,亦即

$$f(u_1,\cdots,u_m) = (u_1,\cdots,u_m,0,\cdots,0),$$

或可理解为:当 $f:M\to N$ 的切映射非退化时,M 就被映为 N 的 m 维子空间,保持其后 $n-m$ 个坐标为零.

图 6.1.12 给出了定理 6.1.13 的图形表示.

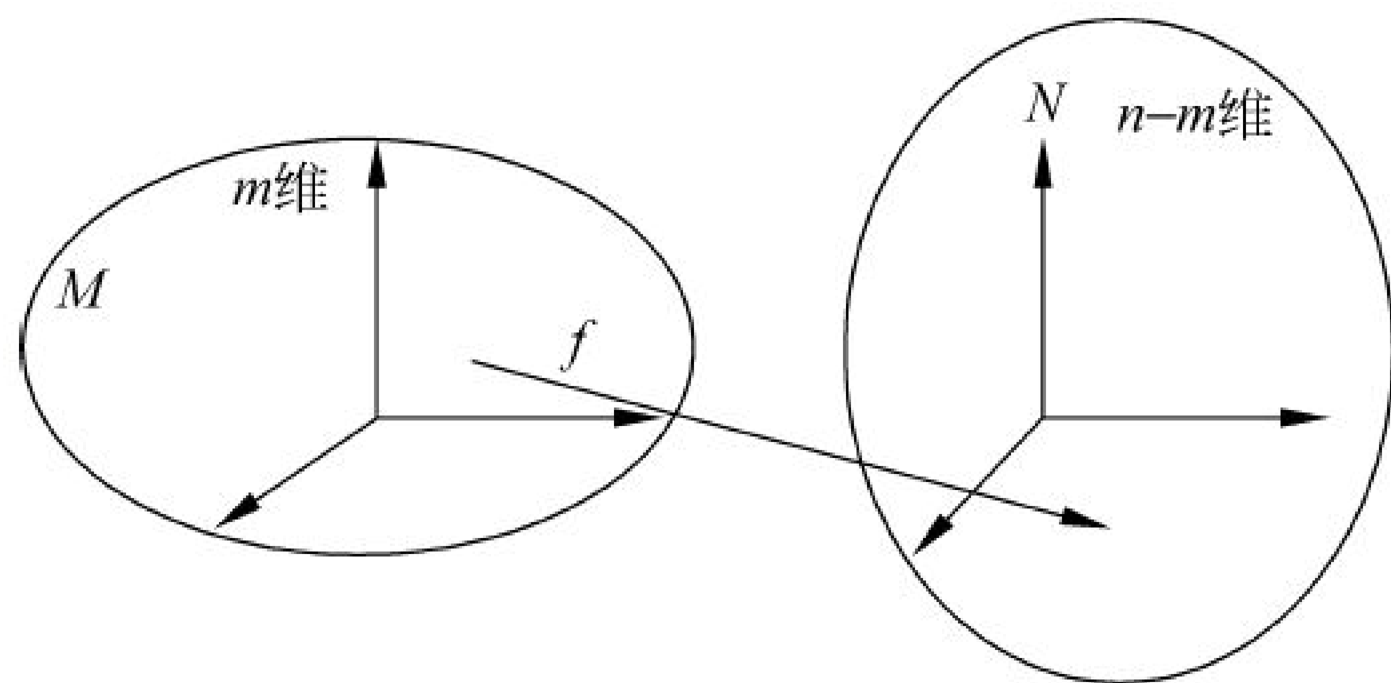


图 6.1.12 定理 6.1.13 的图解

3. 嵌入子流形、浸入子流形

定义 6.1.17(嵌入子流形、浸入子流形) 设$(M,\mathbb{A})$,$(N,\mathbb{B})$是两个光滑流形,若有光滑映射 $\varphi:M\to N$,使得 (1) $\varphi:M\to N$ 是单一的;(2) $\forall\, p\in M$,切映射 $\varphi_*:T_p(M)\to T_{q=\varphi(p)}(N)$都是非退化的,则称$(M,\varphi)\equiv\varphi(M)$是 N 的**嵌入光滑子流形**,称 φ 为 M 到 N 的**嵌入映射**.

若映射 $\varphi:M\to N$ 只满足条件(1),则称 M 为 N 的**浸入子流形**,并称映射 φ 为一个**浸入映射**.

浸入映射只是在局部上是单一的,但不能保证大范围单一;浸入子流形与嵌入子流形的区别在于:像集 $\varphi(M)$是否有自交点.

例 6.1.3(开子流形) 设$(N,\mathbb{B})$是 n 维光滑流形,$U\subset N$ 为开子集.将 N 的光滑流形

结构限制在 U 上,所得到的 U 上的光滑结构,成为与 N 的维数相同的光滑流形.

令 $I:U\to N$ 为恒同映射,则 $I(U)=(U,I)$ 为 N 的一个嵌入子流形,称为 N 的开子流形.

例如,$\mathbb{R}^n\subset\mathbb{R}^{n+1}$ 及恒同映射 $I:\mathbb{R}^n\to\mathbb{R}^{n+1}$ 给出了$\mathbb{R}^{n+1}$的一个 n 维开子流形$\mathbb{R}^n$. 它也是一个嵌入子流形.

例 6.1.4(闭子流形) 设$(N,\mathbb{B})$是 n 维光滑流形,$(M,\mathbb{A})$是 m 维光滑流形,$\varphi:M\to N$ 是光滑映射,于是,$\varphi(M)$是 N 的一个光滑子流形. 如果

(1) $\varphi(M)\subset N$ 是 N 的闭子集;

(2) $\forall q\in\varphi(M)$,存在一个局部坐标系$\{(V,v_k)\}$,使得 $\varphi(M)\cap V$ 是由方程 $v_{m+1}=\cdots=v_n=0$确定,则称 $\varphi(M)\equiv(M,\varphi)$是 N 的一个闭子流形.

例如,单位球面 $S^n\subset\mathbb{R}^{n+1}$及恒同映射 $I:S^n\to\mathbb{R}^{n+1}$给出$\mathbb{R}^{n+1}$的一个 n 维闭子流形 S^n,它也是一个嵌入子流形.

例 6.1.5(浸入子流形) 令 $f:\mathbb{R}^1\to\mathbb{R}^2$ 为映射 $f(t)=\left(2\cos\left(t-\frac{\pi}{2}\right),\sin 2\left(t-\frac{\pi}{2}\right)\right)$,则 $f(\mathbb{R})\subset\mathbb{R}^2$ 是$\mathbb{R}^2$ 的浸入子流形(图 6.1.13(a)),但却不是嵌入的.

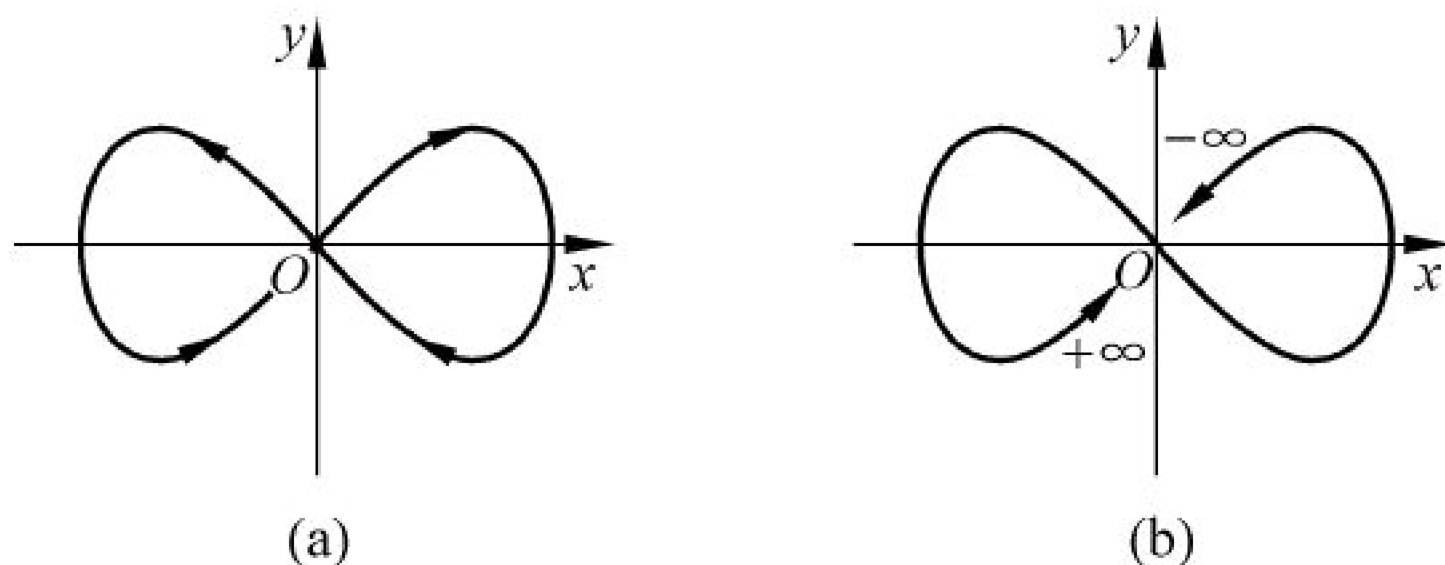


图 6.1.13 例 6.1.5、例 6.1.6

例 6.1.6(嵌入子流形) 令 $g:\mathbb{R}^1\to\mathbb{R}^2$ 为映射

$$g(t)=\left(2\cos\left(2\arctan t+\frac{\pi}{2}\right),\sin 2\left(2\arctan t+\frac{\pi}{2}\right)\right),$$

则 $g(\mathbb{R})$是$\mathbb{R}^2$ 的嵌入子流形(图 6.1.13(b)). 当 $t=0$ 时,$g(0)=(0,0)$,而当 $t\to\pm\infty$时,$g(t)\to(0,0)$.

例 6.1.7(嵌入子流形) 令 $F(t)=\begin{cases}\left(\dfrac{3}{t^2},\sin\pi t\right), & 1\leqslant t<+\infty,\\ f(t), & -1\leqslant t\leqslant 1,\\ (0,t+2), & -\infty<t\leqslant -1,\end{cases}$ 其中 $f(t)$在区间 $t\in[-1,1]$中是连接点$(x,y)=(3,0)$与点$(x,y)=(0,1)$的一条光滑曲线;当 $t\to+\infty$时,曲线无限靠近它自己在 $t\in[-3,-1]$之间的部分(图 6.1.14(a)). 则$F(\mathbb{R})=(\mathbb{R},F)$是$\mathbb{R}$ 到 $\mathbb{R}^2$ 的嵌入子流形.

例 6.1.8(环面) 环面 $T^2=S^1\times S^1$(图 6.1.14(b))可视为**平面上一个单位正方形的两**

组对边"分别等同起来"得到的二维流形. 亦即,环面的点用一对有序实数(x,y)表示,其中$x,y\in\mathbb{R}$,x、y都是模1的实数.

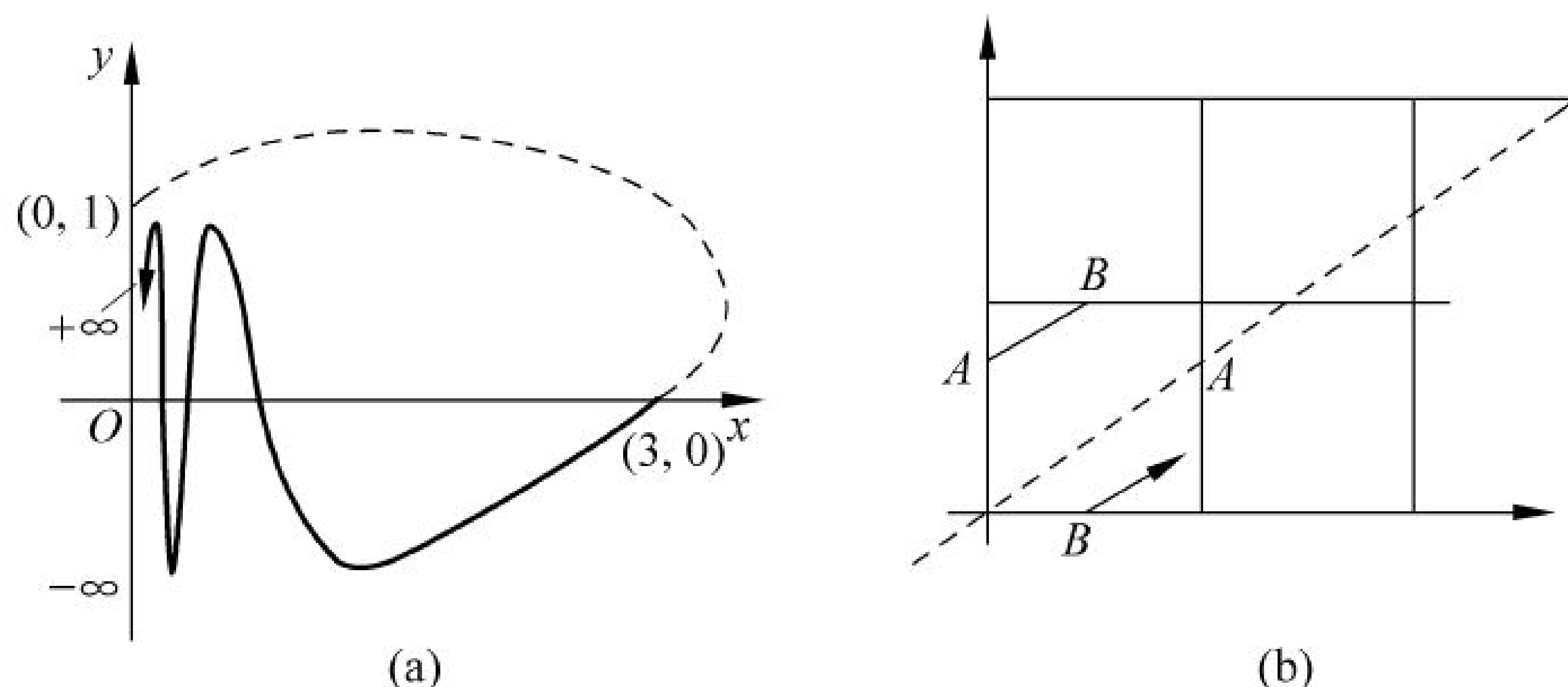


图 6.1.14 例 6.1.7、例 6.1.8

也可以用映射表示. 取两个实数a,b,使得$a:b$是无理数. 考虑映射$\varphi:\mathbb{R}^1\to T^2$,

$$\varphi(t)=(x,y)=(at\,\mathrm{mod}1,bt\,\mathrm{mod}1).$$

显然,$(\mathbb{R},\varphi)\subset T^2$是$T^2$的嵌入子流形,并且像集$\varphi(\mathbb{R})\subset T^2$是$T^2$中处处稠密的子集. 当$a:b$为有理数时,$\varphi(\mathbb{R})\subset T^2$给出环面是$T^2$的一个浸入子流形.

4. 正则子流形

对于光滑流形$(N,\mathbb{B})$的嵌入子流形$\varphi(M)$,要求$\varphi:M\to N$是单一映射,因此可以视子流形$\varphi(M)$的结构与$(M,\mathbb{A})$的结构一致,使得$\varphi:M\to\varphi(M)$是微分同胚. 然而,$\varphi(M)$又是N的子集,因此又可得到一个子拓扑结构,这个结构与$\varphi(M)$从M得到的结构可能相同,也可能不同. 一般来说,后者这个结构比前者要细.

定义 6.1.18(正则子流形) 设$(M,\mathbb{A})$,$(N,\mathbb{B})$是两个光滑流形,若光滑映射$\varphi:M\to N$是M到$\varphi(M)$的同胚映射,则称$\varphi(M)=(M,\varphi)\subset N$是$N$的正则子流形,并称$\varphi:M\to N$为**正则嵌入映射**.

定理 6.1.14(正则子流形的特征性质) 设$(M,\mathbb{A})$为m维光滑流形,$(N,\mathbb{B})$为n维光滑流形,则(M,φ)是N的正则子流形,当且仅当$(M,\varphi)\subset N$是N的一个开子流形的闭子流形.

证 充分性 只要证明N的闭子流形(M,φ)必定是正则子流形即可.

设$(M,\varphi)\subset N$是闭子流形,任取一点$p\in M$,由定义,必定存在点$q=\varphi(p)$在N中的坐标系$\{(V,v_k)\}$,使得$\varphi(M)\cap V$由方程

$$v_{m+1}=\cdots=v_n=0 \tag{6.1.16}$$

确定. 进而,由φ的连续性,存在p的局部坐标系$\{(U,u_j)\}$,使得$\varphi(U)\subset V$. 于是,不妨假设$u_j(p)=0(1\leqslant j\leqslant m)$,$v_k(q)=0(m+1\leqslant k\leqslant n)$,且$V=\{(v_1,\cdots,v_n):|v_k|<\delta\}$,$\delta>0$. 因此,

$\varphi(U)\subset\varphi(M)\cap V$.

下面只要证明 $\varphi^{-1}:\varphi(M)\subset N\to M$ 也是连续映射. 为此,也只要证明:可取适当小的 $\delta_1>0$,使得 $\varphi(M)\cap V_1$ 在 φ 之下的像成立 $\varphi^{-1}(\varphi(M)\cap V_1)\subset U$,其中 V_1 待定.

由(6.1.15)式,映射$\varphi|_U$ 的局部坐标可表示为 $\begin{cases} v_j=\varphi_j(u_1,\cdots,u_m), & 1\leqslant j\leqslant m, \\ v_k=0, & m+1\leqslant k\leqslant n, \end{cases}$ 因此 $\left.\frac{\partial(\varphi_1,\cdots,\varphi_m)}{\partial(u_1,\cdots,u_m)}\right|_{u_j}\neq 0$. 于是,据定理 6.1.11,存在 δ_1,$0<\delta_1<\delta$,使得函数组$(\varphi_1,\cdots,\varphi_m)$有反函数

$$u_j=\psi_j(v_1,\cdots,v_m),\qquad |v_j|<\delta_1.$$

取 $V_1=\{(v_1,\cdots,v_m,v_{m+1},\cdots,v_n):|v_k|<\delta_1\}$,则 $\varphi(M)\cap V_1$ 的逆像集 $\varphi^{-1}(\varphi(M)\cap V_1)$落在 U 中,这就说明 φ^{-1}连续. 充分性得证.

必要性 设(M,φ)是 N 的正则子流形,我们证明:存在 N 的一个开子流形 $W\subset N$,使得(M,φ)是 W 中的闭子流形.

事实上,设 $p\in M$,则对点 p 的任一邻域 $U\subset M$,必存在点 $q=\varphi(p)$在 N 中的一个邻域 V,使得 $\varphi(U)=\varphi(M)\cap V$. 于是,由定理 6.1.13,存在点 p 的局部坐标系$\{(U_1,u_j)\}$与 q 的局部坐标系$\{(V_1,v_k)\}$,使得 $\varphi(U_1)\subset V_1$,并且$\varphi|_{U_1}$ 可用局部坐标表示为

$$\varphi(u_1,\cdots,u_m)=(u_1,\cdots,u_m,0,\cdots,0).\tag{6.1.17}$$

不失一般,可设 $U_1\subset U$,故可取 $V_1\subset V$,因此,$\varphi(U_1)=\varphi(M)\cap V_1$. 再由(6.1.17)式可知,$\varphi(M)\cap V_1$ 是由方程

$$v_{m+1}=\cdots=v_n=0\tag{6.1.18}$$

定义的.

现在来构造开子流形 $W\subset N$. 对每一点 $q=\varphi(p)\in\varphi(M)$,如上构成 $V_q\equiv V_1$,它是点 q 在 N 中的坐标域,使得(6.1.18)式成立. 令 $W=\bigcup\limits_{q\in\varphi(M)}V_q$,显然,$W$ 是 N 的包含 $\varphi(M)$的开子流形.

最后,只要证明 $\varphi(M)$作为 N 的拓扑子空间是 W 中的相对闭子集,亦即证明 $W\cap\overline{\varphi(M)}=\varphi(M)$,其中$\overline{\varphi(M)}$是集合 $\varphi(M)$在 N 中的闭包.

事实上,任取 $s\in W\cap\overline{\varphi(M)}$,据 $W=\bigcup\limits_{q\in\varphi(M)}V_q$,存在 $q\in\varphi(M)$,使得 $s\in V_q$,由(6.1.17)式知,$\varphi(M)\cap V_q$是 V_q中的一个 m 维坐标面,故 $\varphi(M)\cap V_q$是 V_q中的一个相对闭子集.

因 $s\in V_q\cap\overline{\varphi(M)}\subset W\cap\overline{\varphi(M)}$,即 s 属于 $\varphi(M)\cap V_q$ 在 V_q 中的相对闭包,故 $s\in\varphi(M)\cap V_q$,从而,$W\cap\overline{\varphi(M)}\subset\varphi(M)$. 所以有 $W\cap\overline{\varphi(M)}=\varphi(M)$. 这就证明了$(M,\varphi)$是 N 的开子流形 W 中的闭子流形.

推论 子流形(M,φ)是光滑流形$(N,\mathbb{B})$的正则子流形,当且仅当:对于每一点 $p\in M$,存在点 $q=\varphi(p)$在 N 中的局部坐标系$\{(V,v_k)\}$,$v_k(q)=0$,使得 $\varphi(M)\cap V$ 是由 $v_{m+1}=\cdots=$

$v_n=0$ 确定的.

定理 6.1.15 设(M,φ)为光滑流形$(N,\mathbb{B})$的子流形,若$(M,\mathbb{A})$是紧致的,则 $\varphi:M\to N$ 是正则嵌入映射.

证 因 $\varphi(M)$作为 N 的拓扑子空间是 T_2 型的,而 $\varphi:M\to\varphi(M)\subset N$ 是紧致空间 M 到 T_2 型空间 $\varphi(M)$的一对一连续映射,所以 $\varphi:M\to\varphi(M)\subset N$ 是一个同胚,亦即(M,φ)是 N 的正则子流形.

下面再给出正则子流形的一个性质.

定理 6.1.16 设$(M,\mathbb{A})$为 m 维紧致光滑流形,则存在正整数 n,以及 M 到$\mathbb{R}^n$ 的光滑映射 $\varphi:M\to\mathbb{R}^n$,使得(M,φ)是$\mathbb{R}^n$ 的正则子流形.

5. Frobenius 定理

回顾定理 6.1.10,是将光滑流形 M 的一个光滑切矢量场 X 的局部坐标系简化的定理.

局部坐标系的简化定理 设$(M,\mathbb{A})$是光滑流形,X 是 M 上的光滑切矢量场,若在点 $p\in M$ 有 $X_p\neq 0$,则存在点 p 的一个局部坐标系$\{(W,w_j)\}$,使得$X\,|_W=\dfrac{\partial}{\partial w_j}$.

现在的问题是:光滑流形$(M,\mathbb{A})$上有 s 个光滑切矢量场 $X_1,\cdots,X_s$,它们在一个邻域 U 内处处线性无关,那么,是否在 U 中的每一点 p,存在局部坐标系$\{(W,w_j)\}$,使得$X_j\,|_W=\dfrac{\partial}{\partial w_j}$,$1\leqslant j\leqslant s$,同时简化?

我们有下面的结论.

定理 6.1.17 设$(M,\mathbb{A})$是光滑流形,$X_1,\cdots,X_s$ 是M 上的s 个在邻域U 内处处线性无关的光滑切矢量场.在 U 的每一点 p,存在局部坐标系$\{(W,w_j)\}$,使得$X_j\,|_W=\dfrac{\partial}{\partial w_j}$,$1\leqslant j\leqslant s$,当且仅当

$$[X_j,X_k]=0,\quad 1\leqslant j,k\leqslant s.$$

还有比较弱的 Frobenius 条件.

定义 6.1.19(s 维切子空间场) 设$(M,\mathbb{A})$为光滑流形,对于每一个 $p\in M$,指定切空间 T_p 的s 维子空间$L^s(p)$,即 L^s 是M 上s 维切子空间场.若$\forall\, p\in M$,在 p 的一个邻域上存在 s 个处处线性无关的光滑切矢量场 $X_1,\cdots,X_s$,使得$\forall\, q\in U$,子空间 $L^s(q)$是由矢量 $X_1(q),\cdots,X_s(q)$张成的,则称 L^s 是**光滑的 s 维切子空间场**,或称 L^s 是**光滑流形 M 上的 s 维光滑分布**,记为$L^s\,|_U=\{X_1,\cdots,X_s\}$.

定义 6.1.20(Frobenius 条件) 设$(M,\mathbb{A})$为光滑流形,L^s 是 M 上 s 维切子空间场.若在任意坐标域 U 上,当 L^s 由处处线性无关的光滑切矢量场 $X_1,\cdots,X_s$ 张成时,$[X_j,X_k]$,

$1\leqslant j,k\leqslant s$ 都可以表示成 X_j 的线性组合，则称 L^s **满足 Frobenius 条件**.

下面的 Frobenius 定理给出了$L^s|_U=\{X_1,\cdots,X_s\}$能否简化为$L^s|_U=\left\{\dfrac{\partial}{\partial w_1},\cdots,\dfrac{\partial}{\partial w_s}\right\}$的结论.

定理 6.1.18 设$(M,\mathbb{A})$是光滑流形，L^s 是定义在 M 的一个开集 U 上的光滑的 s 维切子空间场，则对于任一点 $p\in U$，存在点 p 的局部坐标系$\{(W,w_j)\}$，$W\subset U$，使得$L^s|_U=\left\{\dfrac{\partial}{\partial w_1},\cdots,\dfrac{\partial}{\partial w_s}\right\}$，当且仅当 L^s 满足 Frobenius 条件.

本章的 6.4.2 节中将看到，“子流形”在“流形上的积分”中起重要中用，为此，我们将其重要概念做一个小结.

光滑流形的子流形分为以下几种：

子流形 设$(M,\mathbb{A})$是 m 维光滑流形，$N\subset M$ 是 M 的一个子集，若对于任一 $p\in N$，存在 p 在M 中的坐标卡(U,φ_U)，使得 $\varphi_U(N\cap U)=(\mathbb{R}^k)\cap\varphi_U(U)$，则称 N 是M 的k 维光滑子流形，简称子流形；

正则子流形 若 N 为M 的k 维光滑子流形，映射 $\Phi:N\to\Phi(N)$，$\Phi(N)\subset M$，是同胚映射，则称(N,Φ)是 M 的正则子流形，并称 Φ 为光滑流形 N 到 M 中的正则嵌入映射；

嵌入子流形 设 N 为M 的k 维光滑子流形，若存在 N 到M 的一个**嵌入映射**，则称 N 是M 的**嵌入子流形**，$h(N)\subset M$；

嵌入映射 称映射 $h:N\to M$ 为 N 到 M 的嵌入映射，若

① $h:N\to M$ 是光滑、一一映射，$p\in N\Rightarrow q=h(p)\in M$，其微分

$$h^*:T_q^*(M)\to T_p^*(N)$$

满足

$$h^*((\mathrm{d}f)_q)=\mathrm{d}(f\circ F),\quad (\mathrm{d}f)_q\in T_q^*(M),$$

其切映射 $h_*:T_p(N)\to T_q(M)$满足

$$\langle h_*X,a\rangle=\langle X,h^*a\rangle,\quad X\in T_p(N),\ a\in T_q^*(M);$$

② $h_*:T_p(N)\to T_q(M)$，在$\forall\, p\in N$ 是非退化的.

浸入子流形 设 N 为M 的k 维光滑子流形，若存在 N 到M 的一个**浸入映射**，则称 N 是M 的**浸入子流形**，$h(N)\subset M$；

浸入映射 称映射 $h:N\to M$ 为 N 到 M 的浸入映射，若仅仅满足上述 ②，亦即，对于 $h:N\to M$ 是光滑映射(但非一一的)，

$$h^*:T_q^*(M)\to T_p^*(N)$$

为其微分，而

$$h_*: T_p(N) \to T_q(M)$$

为其切映射. 若切映射 $h_*: T_p(N) \to T_q(M)$ 在 $\forall p \in N$ 是非退化的,则称 $h: N \to M$ 为**浸入映射**.

正则嵌入子流形 若 N 为 M 的 k 维光滑嵌入子流形,且嵌入映射 $h: N \to M$ 是 N 到 $h(N)$ 的同胚映射,则称 N 为 M 的**正则嵌入子流形**.

6.2 外代数

6.2.1 (r,s) 型张量、(r,s) 型张量空间

在第 2 章中,我们介绍了数域 $\mathbb{F}$(实数域或复数域)上的 n 维线性空间 V,m 维线性空间 W 及其对偶空间 V^*,W^*(V,W 上的线性泛函空间)的张量积. 对于 $v^* \in V^* = L(V,\mathbb{F})$,$w^* \in W^* = L(W,\mathbb{F})$,定义 v^* 与 w^* 的张量积 $v^* \otimes w^*: V \times W \to \mathbb{F}$ 为

$$v^* \otimes w^*(v,w) = v^*(v) \cdot w^*(w) = \langle v, v^* \rangle \cdot \langle w, w^* \rangle, \quad v \in V, w \in W;$$

定义**张量积空间**

$$V^* \otimes W^* \equiv \operatorname{span}\{v^* \otimes w^*: v \in V^*, w^* \in W^*\}. \tag{6.2.1}$$

$\forall f, g \in V^* \otimes W^*$, $\alpha \in \mathbb{F}$,在运算

$$(f+g)(x,y) = f(x,y) + g(x,y), \quad (x,y) \in V \times W;$$
$$(\alpha f)(x,y) = \alpha f(x,y), \quad (x,y) \in V \times W$$

之下,$V^* \otimes W^*$ 成为数域 $\mathbb{F}$ 上的线性空间.

若 $\mathfrak{A} = \{a_1, a_2, \cdots, a_n\}$ 是 n 维线性空间 V 的基,$\mathfrak{A}^* = \{a_j^*, 1 \leqslant j \leqslant n\}$ 是 V^* 的关于基 $\mathfrak{A}$ 的对偶基,满足

$$a_j^*(a_k) = \delta_{kj}, \quad 1 \leqslant j, k \leqslant n; \tag{6.2.2}$$

相应地,$\mathfrak{B} = \{b_1, b_2, \cdots, b_m\}$ 是 m 维线性空间 W 的基,$\mathfrak{B}^* = \{b_k^*, 1 \leqslant k \leqslant m\}$ 是 W^* 的关于基 $\mathfrak{B}$ 的对偶基,满足

$$b_j^*(b_k) = \delta_{kj}, \quad 1 \leqslant j, k \leqslant m. \tag{6.2.3}$$

于是,对于张量积运算

$$\otimes: (v^*, w^*) \in V^* \times W^* \to v^* \otimes w^* \in L(V \times W; \mathbb{F}),$$

由于它的双线性性,故

$$v^* \otimes w^* = \sum_{j,k} (v^*(a_j) w^*(b_k))(a_j^* \otimes b_k^*),$$

因此,$\{a_j^* \otimes b_k^*: 1 \leqslant j \leqslant n, 1 \leqslant k \leqslant m\}$ 是 $V^* \otimes W^*$ 的基,而 $V^* \otimes W^*$ 是 $\mathbb{F}$ 上的 $n \times m$ 维线性空间,并且

$$V^* \otimes W^* \leftrightarrow L(V \times W; \mathbb{F}). \tag{6.2.4}$$

进而,有

$$\langle a_j \otimes b_k, a_l^* \otimes b_s^* \rangle = \delta_{jl}\delta_{ks} = \begin{cases} 1, & (j,k)=(l,s), \\ 0, & (j,k)\neq(l,s), \end{cases}$$

故$\{a_j\otimes b_k: 1\leqslant j\leqslant n, 1\leqslant k\leqslant m\}$是$V\otimes W$的基，$\{a_j^*\otimes b_k^*: 1\leqslant j\leqslant n, 1\leqslant k\leqslant m\}$是$V^*\otimes W^*$的基，并且它们互为对偶基；于是$V^*\otimes W^*=(V\otimes W)^*$. 以这些知识为基础，我们研究$(r,s)$型张量与$(r,s)$型张量空间.

1. (r,s)型张量

定义 6.2.1（**张量及其反变阶数、协变阶数**） 作两个张量积空间$\underbrace{V\otimes\cdots\otimes V}_{r}$与$\underbrace{V^*\otimes\cdots\otimes V^*}_{s}$的张量积

$$V_s^r = \underbrace{V\otimes\cdots\otimes V}_{r}\otimes\underbrace{V^*\otimes\cdots\otimes V^*}_{s}, \tag{6.2.5}$$

其中的元$x\in V_s^r=\underbrace{V\otimes\cdots\otimes V}_{r}\otimes\underbrace{V^*\otimes\cdots\otimes V^*}_{s}$称为$(r,s)$**型张量**，$r$称为$x$的**反变阶数**，$s$称为$x$的**协变阶数**. 特别地，约定$V_0^0=\mathbb{F}$.

$V_0^r=\underbrace{V\otimes\cdots\otimes V}_{r}$中的元称为$r$**阶反变张量**（$r$-**order inverse variant tensor**）；$V_0^1=V$中的元称为**反变张量**（即1阶反变张量）；

$V_s^0=\underbrace{V^*\otimes\cdots\otimes V^*}_{s}$中的元称为$s$**阶协变张量**（$s$-**order covariant tensor**）；$V_1^0=V^*$中的元称为**协变张量**（即1阶协变张量）.

我们有下面结论：

(1) $\dim V_s^r=n^{r+s}$.

(2) $V_s^r \leftrightarrow L(\underbrace{V^*\times\cdots\times V^*}_{r}\times\underbrace{V\times\cdots\times V}_{s};\mathbb{F})$，

因此，(r,s)型张量$x\in V_s^r=\underbrace{V\otimes\cdots\otimes V}_{r}\otimes\underbrace{V^*\otimes\cdots\otimes V^*}_{s}$是定义在积线性空间

$$\underbrace{V^*\times\cdots\times V^*}_{r}\times\underbrace{V\times\cdots\times V}_{s}$$

上的$\mathbb{F}$值$r+s$**重线性函数**（$r+s$**重线性泛函**）.

(3) 对于n维线性空间V的基$\mathfrak{A}=\{a_j, 1\leqslant j\leqslant n\}$，对偶空间$V^*$的对偶基为$\mathfrak{A}^*=\{a_j^*, 1\leqslant j\leqslant n\}$，则张量空间$V_s^r=\underbrace{V\otimes\cdots\otimes V}_{r}\otimes\underbrace{V^*\otimes\cdots\otimes V^*}_{s}$的基是

$$\{a_{j_1}\otimes\cdots\otimes a_{j_r}\otimes a_{k_1}^*\cdots\otimes a_{k_s}^*, 1\leqslant j_1,\cdots,j_r,k_1,\cdots,k_s\leqslant n\},$$

而(r,s)型张量$x\in V_s^r$可惟一地表示为

$$x=\sum_{1\leqslant j_1,\cdots,j_r,k_1,\cdots,k_s\leqslant n} x_{k_1,\cdots,k_s}^{j_1,\cdots,j_r}(a_{j_1}\otimes\cdots\otimes a_{j_r})\otimes(a_{k_1}^*\otimes\cdots\otimes a_{k_s}^*),$$

并且有

$$x_{k_1,\cdots,k_s}^{j_1,\cdots,j_r}=x(a_{j_1},\cdots,a_{j_r},a_{k_1}^*,\cdots,a_{k_s}^*)=\langle a_{j_1}\otimes\cdots\otimes a_{j_r}\otimes a_{k_1}^*\otimes\cdots\otimes a_{k_s}^*,x\rangle.$$

记住 **Einstein 和式约定**：将和式

$$x=\sum_{1\leqslant j_1,\cdots,j_r,k_1,\cdots,k_s\leqslant n} x_{k_1,\cdots,k_s}^{j_1,\cdots,j_r}(a_{j_1}\otimes\cdots\otimes a_{j_r})\otimes(a_{k_1}^*\otimes\cdots\otimes a_{k_s}^*)$$

简记为

$$x=x_{k_1,\cdots,k_s}^{j_1,\cdots,j_r}(a_{j_1}\otimes\cdots\otimes a_{j_r})\otimes(a_{k_1}^*\otimes\cdots\otimes a_{k_s}^*),\tag{6.2.6}$$

或 $x=x_{k_1\cdots k_s}^{j_1\cdots j_r}$.

2. (r,s)型张量空间

定义 6.2.2（(r,s)型张量空间及其上的运算） 记(r,s)型张量的全体为

$$V_s^r=\underbrace{V\otimes\cdots\otimes V}_{r}\otimes\underbrace{V^*\otimes\cdots\otimes V^*}_{s}.$$

同类型的张量可以做加法、数乘，即任取 $x,y\in V_s^r$，$\alpha\in\mathbb{F}$，若

$$x=x_{k_1,\cdots,k_s}^{j_1,\cdots,j_r}(a_{j_1}\otimes\cdots\otimes a_{j_r})\otimes(a_{k_1}^*\otimes\cdots\otimes a_{k_s}^*),$$

$$y=y_{k_1,\cdots,k_s}^{j_1,\cdots,j_r}(a_{j_1}\otimes\cdots\otimes a_{j_r})\otimes(a_{k_1}^*\otimes\cdots\otimes a_{k_s}^*),$$

则

$$x+y=(x_{k_1,\cdots,k_s}^{j_1,\cdots,j_r}+y_{k_1,\cdots,k_s}^{j_1,\cdots,j_r})(a_{j_1}\otimes\cdots\otimes a_{j_r})\otimes(a_{k_1}^*\otimes\cdots\otimes a_{k_s}^*),\tag{6.2.7}$$

$$\alpha x=(\alpha x_{k_1,\cdots,k_s}^{j_1,\cdots,j_r})(a_{j_1}\otimes\cdots\otimes a_{j_r})\otimes(a_{k_1}^*\otimes\cdots\otimes a_{k_s}^*),\tag{6.2.8}$$

从而 $V_s^r=\underbrace{V\otimes\cdots\otimes V}_{r}\otimes\underbrace{V^*\otimes\cdots\otimes V^*}_{s}$ 成为$\mathbb{F}$上的线性空间，称为(r,s)型张量线性空间，简称**(r,s)型张量空间.**

在 V_s^r 上可定义张量积运算如下：

定义 6.2.3（张量积） 设 $x\in V_{s_1}^{r_1}$ 为(r_1,s_1)型张量，$y\in V_{s_2}^{r_2}$ 为(r_2,s_2)型张量，则 x 与 y 的张量积定义为

$$\begin{aligned}&x\otimes y(v_1^*,\cdots,v_{r_1+r_2}^*,v_1,\cdots,v_{s_1+s_2})\\&=x(v_1^*,\cdots,v_{r_1}^*,v_1,\cdots,v_{s_1})\cdot y(v_{r_1+1}^*,\cdots,v_{r_1+r_2}^*,v_{s_1+1},\cdots,v_{s_1+s_2}).\end{aligned}\tag{6.2.9}$$

在基取定后，$x\otimes y$ 的分量是 x 与 y 的分量之积，即

$$(x\otimes y)_{k_1,\cdots,k_{s_1+s_2}}^{j_1,\cdots,j_{r_1+r_2}}=x_{k_1,\cdots,k_{s_1}}^{j_1,\cdots,j_{r_1}}\cdot y_{k_{s_1+1},\cdots,k_{s_1+s_2}}^{j_{r_1+1},\cdots,j_{r_1+r_2}}.\tag{6.2.10}$$

张量积运算适合分配律与结合律.

定义 6.2.4((r,0)型与(0,s)型张量空间) 空间 $T^r(V)\equiv V_0^r=\underbrace{V\otimes\cdots\otimes V}_{r}$ 是一个(r,0)型张量空间,加法、数乘可如下定义:

设 $x,y\in T^r(V)$,且在选定的坐标下,有

$$x=x^{j_1,\cdots,j_r}(a_{j_1}\otimes\cdots\otimes a_{j_r}),\quad y=y^{j_1,\cdots,j_r}(a_{j_1}\otimes\cdots\otimes a_{j_r}),$$

则

$$x+y=(x^{j_1,\cdots,j_r}+y^{j_1,\cdots,j_r})(a_{j_1}\otimes\cdots\otimes a_{j_r}),$$

$$\alpha x=(\alpha x^{j_1,\cdots,j_r})(a_{j_1}\otimes\cdots\otimes a_{j_r}).$$

同理,可定义(0,s)**型张量空间** $T^s(V^*)=(V^*)_s^0=\underbrace{V^*\otimes\cdots\otimes V^*}_{s}$,及其中的运算.

下面是(r,0)型张量空间与(0,r)型张量空间的对偶性定理.

定理 6.2.1 (r,0)型张量空间 $T^r(V)$ 与(0,r)型张量空间 $T^r(V^*)$ 是彼此对偶的,亦即 $(T^r(V))^*=T^r(V^*)$,且 $\dim T^r(V)=\dim T^r(V^*)=n^r$.

定义 6.2.5(张量空间中的配合) 在张量空间 $T^r(V)$ 与 $T^r(V^*)$ 的元之间,定义配合

$$\langle v_1\otimes\cdots\otimes v_r,v_1^*\otimes\cdots\otimes v_r^*\rangle=\langle v_1,v_1^*\rangle\cdots\langle v_r,v_r^*\rangle,\ v_j\in V,v_k^*\in V^*,1\leqslant j,k\leqslant n.$$

"配合"是张量空间 $T^r(V)$ 中的元 $v_1\otimes\cdots\otimes v_r$ 对 $T^r(V^*)$ 中的元 $v_1^*\otimes\cdots\otimes v_r^*$ 的作用.

3. 张量空间 $T^r(V)$ 上的自同态 $\sigma\in\mathfrak{T}(r)$,r 次置换群 $\mathfrak{T}(r)$

回顾1.2节中定义的置换群:对于 r 个元的集合 $S=\{1,2,\cdots,r\}$,称 S 到自身的一一映射 $\sigma:S\to S$ 为一个置换,使得

$$j\in S\to\sigma(j)\in S.$$

记 S 上所有置换所成的集为

$$\mathfrak{T}(r)=\{\sigma:\sigma\{1,2,\cdots,r\}=\{\sigma(1),\sigma(2),\cdots,\sigma(r)\}\},$$

记 σ 为

$$\sigma=\begin{pmatrix}1&2&\cdots&r\\\sigma(1)&\sigma(2)&\cdots&\sigma(r)\end{pmatrix},$$

显然,$\mathfrak{T}(r)$ 中有 $r!$ 个元.

定义 $\mathfrak{T}(r)$ 中的元的运算 $\sigma,\tau\in\mathfrak{T}(r)\Rightarrow\tau\circ\sigma$ 为两个置换的复合,即

$$(\tau\circ\sigma)(s)=\tau(\sigma(s)),\quad\forall s\in S.$$

于是,在复合运算 $\tau\circ\sigma$ 之下,$(\mathfrak{T}(r),\circ)$ 成为一个群,其单位元为 $\sigma(s)=s,s\in S$,并且有 $\sigma\in\mathfrak{T}(r)\Rightarrow\tau^{-1}\in\mathfrak{T}(r)$.

$\forall\sigma\in\mathfrak{T}(r)$,若 $\sigma=\begin{pmatrix}1&2&\cdots&r\\\sigma(1)&\sigma(2)&\cdots&\sigma(r)\end{pmatrix}$,对于 $k<j$,做 $\sigma(j)-\sigma(k)$,并让 k 与 j 跑遍 S,再作乘积

$$\prod_{1\leqslant k<j\leqslant r}(\sigma(j)-\sigma(k))=\pm(2!)(3!)\cdots((n-1)!),\tag{6.2.11}$$

若此积为正数，则称 σ 为偶置换，若此积为负数，则称 σ 为奇置换.

将置换概念用于张量空间，由于

$$\sigma\{1,2,\cdots,r\}=\{\sigma(1),\sigma(2),\cdots,\sigma(r)\}=\{j_1,j_2,\cdots,j_r\},\quad 1\leqslant j_1,j_2,\cdots,j_r\leqslant r,$$

$\forall\sigma\in\mathfrak{T}r)$ 决定了张量空间 $T^r(V)$ 的一个自同态 σ: $T^r(V)\to T^r(V)$.

对于 $x\in T^r(V)=V_0^r=\underbrace{V\otimes\cdots\otimes V}_{r}$，定义

$$\sigma(x(v_1^*,\cdots,v_r^*))=x(v_{\sigma(1)}^*,\cdots,v_{\sigma(r)}^*),\quad v_j^*\in V^*.\tag{6.2.12}$$

注 1 (6.2.12)式的意义是：因 $T^r(V)$ 是 $T^r(V^*)$ 的对偶空间，$(T^r(V))^*=T^r(V^*)$，故

$$\forall x\in T^r(V)\Rightarrow x=x^{j_1,\cdots,j_r}(a_{j_1}\otimes\cdots\otimes a_{j_r})$$

$$\Rightarrow x(v_1^*,\cdots,v_r^*)\in\mathbb{F},\quad\forall(v_1^*,\cdots,v_r^*)\in T^r(V^*).$$

因此，$\forall\sigma\in\mathfrak{T}r)$，$\sigma$: $T^r(V)\to T^r(V)$ 将 $x\in T^r(V)$ 映为 $\sigma x\in T^r(V)$，满足

$$(\sigma x)(v_1^*,\cdots,v_r^*)=\sigma(x(v_1^*,\cdots,v_r^*))=x(v_{\sigma(1)}^*,\cdots,v_{\sigma(r)}^*),\quad v_j^*\in V^*.$$

注 2 自同态 σ: $T^r(V)\to T^r(V)$ 满足

$$\sigma(x+y)=\sigma(x)+\sigma(y);\quad\sigma(\alpha x)=\alpha\sigma(x),\ \alpha\in\mathbb{F}.$$

定理 6.2.2 若 $x\in T^r(V)$，$x=v_1\otimes\cdots\otimes v_r$，则

$$\sigma x=v_{\sigma^{-1}(1)}\otimes\cdots\otimes v_{\sigma^{-1}(r)},\quad\forall\sigma\in\mathfrak{T}r),$$

其中 σ^{-1} 是 σ 的逆元.

4. $T^r(V)$ 张量空间中的对称反变张量、反对称反变张量

定义 6.2.6(对称反变张量、反对称反变张量) 设 $x\in T^r(V)$，若对于任意的 $\sigma\in\mathfrak{T}r)$，都有

$$\sigma x=x,\tag{6.2.13}$$

则称 x 是**对称的 r 阶反变张量(symmetric r-order inverse variant tensor)**；若对于任意的 $\sigma\in\mathfrak{T}r)$，都有

$$\sigma x=(\mathrm{sgn}\sigma)\cdot x,\tag{6.2.14}$$

则称 x 是**反对称的 r 阶反变张量(skew-symmetric r-order inverse variant tensor)**；其中 $\mathrm{sgn}\sigma$ 表示置换 σ 的符号，即

$$\mathrm{sgn}\sigma=\begin{cases}1, & \sigma\text{ 是偶置换},\\-1, & \sigma\text{ 是奇置换}.\end{cases}$$

定理 6.2.3 若 $x\in T^r(V)$，则

(1) x 是对称反变张量，当且仅当 x 的分量关于各指标是对称的；

(2) x 是反对称反变张量，当且仅当 x 的分量关于各指标是反对称的.

证 设 V 的基为 $\{a_1,\cdots,a_n\}$，V^* 的对偶基为 $\{a_1^*,\cdots,a_n^*\}$.

(1) 当 x 是对称反变张量时，对任意 $\sigma\in\mathfrak{T}r)$，有

$$x^{j_1,\cdots,j_r}(a_{j_1}^*,\cdots,a_{j_r}^*)=x(a_{j_1}^*,\cdots,a_{j_r}^*)=\sigma x(a_{j_1}^*,\cdots,a_{j_r}^*)$$
$$=x(a_{j_{\sigma(1)}}^*,\cdots,a_{j_{\sigma(r)}}^*)=x^{j_{\sigma(1)},\cdots,j_{\sigma(r)}}(a_{j_1}^*,\cdots,a_{j_r}^*),$$

反之亦然.

(2) 当 x 是对称反变张量时，对任意 $\sigma\in\mathfrak{T}r)$，有

$$x^{j_1,\cdots,j_r}(a_{j_1}^*,\cdots,a_{j_r}^*)=x(a_{j_1}^*,\cdots,a_{j_r}^*)=(\mathrm{sgn}\sigma)\cdot\sigma x(a_{j_1}^*,\cdots,a_{j_r}^*)$$
$$=(\mathrm{sgn}\sigma)\cdot x(a_{j_{\sigma(1)}}^*,\cdots,a_{j_{\sigma(r)}}^*)=(\mathrm{sgn}\sigma)\cdot x^{j_{\sigma(1)},\cdots,j_{\sigma(r)}}(a_{j_1}^*,\cdots,a_{j_r}^*),$$

反之亦然. 定理得证.

$T^r(V)$ 中**对称 r 阶反变张量**的全体，记为 $P^r(V)$；$T^r(V)$ 中**反对称 r 阶反变张量**的全体，记为 $\Lambda^r(V)$.

由于置换 $\sigma\in\mathfrak{T}r)$ 是 $T^r(V)$ 上的自同态，所以对称张量的和仍然是对称的，反对称张量的和仍然是反对称的，故 $P^r(V)$，$\Lambda^r(V)$ 都是 $T^r(V)$ 的线性子空间.

定义 6.2.7(对称化算子、反对称化算子) 对于任意 $x\in T^r(V)$，令

$$S_r(x)=\frac{1}{r!}\sum_{\sigma\in\mathfrak{T}r)}\sigma x,\quad A_r(x)=\frac{1}{r!}\sum_{\sigma\in\mathfrak{T}r)}(\mathrm{sgn}\sigma)\cdot\sigma x.\tag{6.2.15}$$

则 $S_r(x),A_r(x)\in T^r(V)$，并且 $S_r,A_r:T^r(V)\to T^r(V)$. 于是算子 S_r,A_r 都是 $T^r(V)$ 上的自同态，称 S_r 为 $T^r(V)$ 上的 **r 阶反变张量的对称化算子(symmetric operator of r-order inverse variant tensor)**，而称 A_r 为 $T^r(V)$ 上的 **r 阶反变张量的反对称化算子(skew-symmetric operator of r-order inverse variant tensor)**.

定理 6.2.4 $P^r(V)=S_r(T^r(V))$，$\Lambda^r(V)=A_r(T^r(V))$.

证 首先证明，张量 $x\in T^r(V)$ 对称化的结果是对称张量，而 $x\in T^r(V)$ 反对称化的结果是反对称张量. 因为 $\forall x\in T^r(V)$，$\forall\tau\in\mathfrak{T}r)$，有

$$\tau(S_r(x))=\frac{1}{r!}\sum_{\sigma\in\mathfrak{T}r)}\tau(\sigma(x))=S_r(x),$$

$$\tau(A_r(x))=\frac{1}{r!}\sum_{\sigma\in\mathfrak{T}r)}(\mathrm{sgn}\sigma)\cdot\tau(\sigma(x))$$
$$=(\mathrm{sgn}\tau)\cdot\frac{1}{r!}\sum_{\sigma\in\mathfrak{T}r)}(\mathrm{sgn}(\tau\circ\sigma))\cdot(\tau\circ\sigma)(x)=(\mathrm{sgn}\tau)\cdot A_r(x).$$

于是，$S_r(T^r(V))\subset P^r(V)$，$A_r(T^r(V))\subset\Lambda^r(V)$.

其次证明，对称张量在对称算子作用下不变，反对称张量在反对称算子作用下不变(证明留作练习). 因此，$P^r(V)\subset S_r(P^r(V))$，$\Lambda^r(V)\subset A_r(\Lambda^r(V))$.

最后得到

$$P^r(V)=S_r(T^r(V)),\quad\Lambda^r(V)=A_r(T^r(V)).\tag{6.2.16}$$

以上讨论全部可用于协变张量.

$T^r(V^*)$中**对称 r 阶协变张量**的全体,记为 $P^r(V^*)$; $T^r(V^*)$中**反对称 r 阶协变张量**的全体,记为 $\Lambda^r(V^*)$.

反对称 r 阶反变张量空间 $\Lambda^r(V)$也称为 V **上的 r 次外矢量空间**,其中的元也称为 **r 次外矢量**,约定 $\Lambda^0(V)=\mathbb{F}$,$\Lambda^1(V)=V$.

反对称 r 阶协变张量空间 $\Lambda^r(V^*)$也称为 V **上的 r 次外形式空间**,其中的元也称为 **r 次外形式**,约定 $\Lambda^0(V^*)=\mathbb{F}$,$\Lambda^1(V^*)=V^*$.

反对称张量在外微分形式的研究中占有极其重要的地位. 其重要性在于:反对称张量可以定义外积运算与外微分形式.

外微分形式是由 E. Cartan 发展起来的,这使得微积分发展到一个新的阶段,使数学研究达到一个新的高度.

6.2.2 张量代数

1. 张量空间 $T^r(V)=V_0^r$ 的分层和

设 $T^r(V)=V_0^r=\underbrace{V\otimes\cdots\otimes V}_{r}$,考虑一种特殊的和

$$T(V)=\sum_{r\geqslant 0}T^r(V),\tag{6.2.17}$$

称为**分层和**(graded sum),其元素 $x\in T(V)$可表示为直和形式

$$x=\sum_{r\geqslant 0}x^r,\quad x^r\in T^r(V),\tag{6.2.18}$$

并且和式中除有限多项外,其余各项都是零.

在分层和 $T(V)$中,可定义元素的加法$+$与数乘 $\alpha\cdot$,使得$(T(V),+,\alpha\cdot)$成为一个线性空间(但它是无穷维线性空间).

2. $T(V)=\sum_{r\geqslant 0}T^r(V)$ 中的乘法,张量代数$(T(V),+,\alpha\cdot,\otimes)$

张量积(即张量的乘法)可以扩大到分层和空间 $(T(V),+,\alpha\cdot)=\left(\sum_{r\geqslant 0}T^r(V),+,\alpha\cdot\right)$ 中进行. 对于 $x,y\in T(V)=\sum_{r\geqslant 0}T^r(V)$,

① 若 $x\in T^{r_1}(V),y\in T^{r_2}(V),r_1\geqslant 0,r_2\geqslant 0$,则定义

$$x\otimes y(v_1^*,\cdots,v_{r_1+r_2}^*)=x(v_1^*,\cdots,v_{r_1}^*)\cdot y(v_{r_1+1}^*,\cdots,v_{r_1+r_2}^*).$$

在取定基之后,$x\otimes y$ 的分量是 x 与 y 的分量之积,有

$$(x\otimes y)^{j_1,\cdots,j_{r_1+r_2}}=x^{j_1,\cdots,j_{r_1}}\cdot y^{j_{r_1+1},\cdots,j_{r_1+r_2}};\tag{6.2.19}$$

显然有 $x\in T^{r_1}(V),y\in T^{r_2}(V)\Rightarrow x\otimes y\in T^{r_1+r_2}(V)$,运算封闭性显然成立.

② 若 $x=\sum_{r\geqslant 0}x_r$，$y=\sum_{r\geqslant 0}y_r$，因两者均为有限和，则可定义

$$x\otimes y=\left(\sum_{r\geqslant 0}x_r\right)\otimes\left(\sum_{r\geqslant 0}y_r\right),\tag{6.2.20}$$

按分配律计算.

定义 6.2.8(张量代数) 张量空间 $T(V)$在运算$+$,$\alpha\cdot$,$\otimes$之下成为一个"代数"，记为

$$(T(V),+,\alpha\cdot,\otimes)=\left(\sum_{r\geqslant 0}T^r(V),+,\alpha\cdot,\otimes\right),$$

称为**线性空间 V 上的张量代数(tensor algebra on linear space V)**，也称为分层代数(graded algebra).

同理，可定义张量空间 V^* 上的张量代数$(T(V^*),+,\alpha\cdot,\otimes)=\left(\sum_{r\geqslant 0}T^r(V^*),+,\alpha\cdot,\otimes\right)$.

注意到，分层和 $T(V)=\sum_{r\geqslant 0}T^r(V)=\sum_{r\geqslant 0}V_0^r$ 与分层和 $T(V^*)=\sum_{r\geqslant 0}T^r(V^*)=\sum_{r\geqslant 0}V_r^0$ 关于运算$\otimes$构成"代数"，但是 $T^r(V)$，$T^r(V^*)$关于运算$\otimes$却不能构成"代数"，因为运算$\otimes$不封闭. 下面引进"外积"，使之成为一个外代数.

3. 外积

1) 反对称 r 阶反变张量空间 $\Lambda^r(V)$中的外积运算

定义 6.2.9(外积运算) 设 $\xi\in\Lambda^k(V)$，$\eta\in\Lambda^l(V)$，令

$$\xi?\eta=A_{k+l}(\xi\otimes\eta),$$

其中 A_{k+l}是反对称化算子，$A_r(\xi)=\dfrac{1}{r!}\sum_{\sigma\in\mathfrak{T}r)}(\operatorname{sgn}\sigma)\cdot\sigma\xi$，则 $\xi?\eta$ 是一个 $k+l$ 次外矢量，称 $\xi?\eta=A_{k+l}(\xi\otimes\eta)$为 ξ 与 η 的**外积(outer product)**.

定理 6.2.5 外积有如下运算规则：对于 $\xi,\xi_1,\xi_2\in\Lambda^k(V)$，$\eta,\eta_1,\eta_2\in\Lambda^l(V)$ 与 $\zeta\in\Lambda^h(V)$，有

(1) **结合律** $(\xi?\eta)?\zeta=\xi?(\eta?\zeta)$；

(2) **分配律** $(\xi_1+\xi_2)?\eta=\xi_1?\eta+\xi_2?\eta$， $\xi?(\eta_1+\eta_2)=\xi?\eta_1+\xi?\eta_2$；

(3) **反交换律** $\xi?\eta=(-1)^{kl}\eta?\xi$.

证 首先，由张量积的线性性质与反对称化算子的线性性质，分配律显然成立.

其次证反交换律. 对于 k 次外矢量 $\xi\in\Lambda^k(V)$，l 次外矢量 $\eta\in\Lambda^l(V)$，由 $\xi?\eta=A_{k+l}(\xi\otimes\eta)\in\Lambda^{k+l}(V)$是反对称张量，所以对于任意置换 $\tau\in\mathfrak{T}k+l)$，有 $\tau(\xi?\eta)=\operatorname{sgn}\tau\cdot(\xi?\eta)$. 取

$$\tau=\begin{pmatrix}1 & \cdots & k & k+1 & \cdots & k+l\\ 1+l & \cdots & k+l & 1 & \cdots & l\end{pmatrix},$$

则 $\operatorname{sgn}\tau=(-1)^{kl}$，所以 $\forall\, v_1^*,\cdots,v_k^*,v_{k+1}^*,\cdots,v_{k+l}^*\in V^*$，得到

$$
\begin{aligned}
\xi ? \eta(v_1^*,\cdots,v_{k+l}^*) &= (-1)^{kl}\xi ? \eta(v_{\tau(1)}^*,\cdots,v_{\tau(k+l)}^*) \\
&= \frac{(-1)^{kl}}{(k+l)!}\sum_{\sigma\in\mathfrak{S}_{k+l}}\operatorname{sgn}\sigma\cdot\sigma(\xi\otimes\eta)(v_{\tau(1)}^*,\cdots,v_{\tau(k+l)}^*) \\
&= \frac{(-1)^{kl}}{(k+l)!}\sum_{\sigma\in\mathfrak{S}_{k+l}}\operatorname{sgn}\sigma\cdot\xi(v_{\sigma\circ\tau(1)}^*,\cdots,v_{\sigma\circ\tau(k)}^*)\cdot\eta(v_{\sigma\circ\tau(k+1)}^*,\cdots,v_{\sigma\circ\tau(k+l)}^*) \\
&= \frac{(-1)^{kl}}{(k+l)!}\sum_{\sigma\in\mathfrak{S}_{k+l}}\operatorname{sgn}\sigma\cdot\eta(v_{\sigma(1)}^*,\cdots,v_{\sigma(l)}^*)\cdot\xi(v_{\sigma(l+1)}^*,\cdots,v_{\sigma(l+k)}^*) \\
&= \frac{(-1)^{kl}}{(k+l)!}\sum_{\sigma\in\mathfrak{S}_{k+l}}\operatorname{sgn}\sigma\cdot\sigma(\eta\otimes\xi)(v_1^*,\cdots,v_l^*,v_{l+1}^*,\cdots,v_{l+k}^*) \\
&= (-1)^{kl}A_{l+k}(\eta\otimes\xi)(v_1^*,\cdots,v_{l+k}^*) = (-1)^{kl}\eta ? \xi(v_1^*,\cdots,v_{l+k}^*).
\end{aligned}
$$

最后证明结合律．对于 $\xi\in\Lambda^k(V)$，$\eta\in\Lambda^l(V)$，$\zeta\in\Lambda^h(V)$，取

$$
v_1^*,\cdots,v_k^*,v_{k+1}^*,\cdots,v_{k+l}^*,v_{k+l+1}^*,\cdots,v_{k+l+h}^*\in V^*,
$$

则

$$
\begin{aligned}
&((\xi ? \eta) ? \zeta)(v_1^*,\cdots,v_k^*,v_{k+1}^*,\cdots,v_{k+l}^*,v_{k+l+1}^*,\cdots,v_{k+l+h}^*) \\
&= A_{(k+l)+h}((\xi ? \eta)\otimes\varsigma)(v_1^*,\cdots,v_k^*,v_{k+1}^*,\cdots,v_{k+l}^*,v_{k+l+1}^*,\cdots,v_{k+l+h}^*) \\
&= \frac{1}{((k+l)+h)!}\sum_{\sigma\in\mathfrak{S}_{k+l+h}}(\operatorname{sgn}\sigma) \\
&\qquad\cdot\sigma((\xi ? \eta)\otimes\varsigma)(v_1^*,\cdots,v_k^*,v_{k+1}^*,\cdots,v_{k+l}^*,v_{k+l+1}^*,\cdots,v_{k+l+h}^*) \\
&= \frac{1}{((k+l)+h)!}\sum_{\sigma\in\mathfrak{S}_{k+l+h}}\operatorname{sgn}\sigma\cdot(\xi ? \eta)(v_{\sigma(1)}^*,\cdots,v_{\sigma(k)}^*,v_{\sigma(k+1)}^*,\cdots,v_{\sigma(k+l)}^*) \\
&\qquad\cdot\varsigma(v_{\sigma(k+l+1)}^*,\cdots,v_{\sigma(k+l+h)}^*) \\
&= \frac{1}{((k+l)+h)!}\sum_{\sigma\in\mathfrak{S}_{k+l+h}}\operatorname{sgn}\sigma\cdot\frac{1}{(k+l)!}\sum_{\tau\in\mathfrak{S}_{k+l}}\operatorname{sgn}\tau\cdot\xi(v_{\sigma\circ\tau(1)}^*,\cdots,v_{\sigma\circ\tau(k)}^*) \\
&\qquad\cdot\eta(v_{\sigma\circ\tau(k+1)}^*,\cdots,v_{\sigma\circ\tau(k+l)}^*)\cdot\varsigma(v_{\sigma\circ\tau(k+l+1)}^*,\cdots,v_{\sigma\circ\tau(k+l+h)}^*) \\
&= \frac{1}{(k+l+h)!}\cdot\frac{1}{(k+l)!}\sum_{\sigma\in\mathfrak{S}_{k+l+h}}\sum_{\tau\in\mathfrak{S}_{k+l}}\operatorname{sgn}\sigma\cdot\operatorname{sgn}\tau\cdot\xi(v_{\sigma\circ\tau(1)}^*,\cdots,v_{\sigma\circ\tau(k)}^*) \\
&\qquad\qquad\cdot\eta(v_{\sigma\circ\tau(k+1)}^*,\cdots,v_{\sigma\circ\tau(k+l)}^*)\cdot\varsigma(v_{\sigma\circ\tau(k+l+1)}^*,\cdots,v_{\sigma\circ\tau(k+l+h)}^*) \\
&= A_{k+l+h}(\xi\otimes\eta\otimes\varsigma)(v_1^*,\cdots,v_k^*,v_{k+1}^*,\cdots,v_{k+l}^*,v_{k+l+1}^*,\cdots,v_{k+l+h}^*);
\end{aligned}
$$

且

$$
\begin{aligned}
&(\xi ? (\eta ? \varsigma))(v_1^*,\cdots,v_k^*,v_{k+1}^*,\cdots,v_{k+l}^*,v_{k+l+1}^*,\cdots,v_{k+l+h}^*) \\
&= A_{k+(l+h)}(\xi\otimes(\eta ? \varsigma))(v_1^*,\cdots,v_k^*,v_{k+1}^*,\cdots,v_{k+l}^*,v_{k+l+1}^*,\cdots,v_{k+l+h}^*) \\
&= \frac{1}{(k+(l+h))!}\sum_{\sigma\in\mathfrak{S}_{k+l+h}}(\operatorname{sgn}\sigma) \\
&\qquad\cdot\sigma(\xi\otimes(\eta ? \varsigma))(v_1^*,\cdots,v_k^*,v_{k+1}^*,\cdots,v_{k+l}^*,v_{k+l+1}^*,\cdots,v_{k+l+h}^*)
\end{aligned}
$$

$$= \frac{1}{(k+l+h)!} \sum_{\sigma \in \mathfrak{S}(k+l+h)} \operatorname{sgn}\sigma \cdot \xi(v^*_{\sigma(1)}, \cdots, v^*_{\sigma(k)})$$

$$\cdot (\eta ? \varsigma)(v^*_{\sigma(k+1)}, \cdots, v^*_{\sigma(k+l)} v^*_{\sigma(k+l+1)}, \cdots, v^*_{\sigma(k+l+h)})$$

$$= \frac{1}{((k+l)+h)!} \sum_{\sigma \in \mathfrak{S}(k+l+h)} \operatorname{sgn}\sigma \cdot \frac{1}{(k+l)!} \sum_{\tau \in \mathfrak{S}(k+l)} \operatorname{sgn}\tau \cdot \xi(v^*_{\sigma\circ\tau(1)}, \cdots, v^*_{\sigma\circ\tau(k)})$$

$$\cdot \eta(v^*_{\sigma\circ\tau(k+1)}, \cdots, v^*_{\sigma\circ\tau(k+l)}) \cdot \varsigma(v^*_{\sigma\circ\tau(k+l+1)}, \cdots, v^*_{\sigma\circ\tau(k+l+h)})$$

$$= \frac{1}{((k+l)+h)!} \cdot \frac{1}{(k+l)!} \sum_{\sigma \in \mathfrak{S}(k+l+h)} \sum_{\tau \in \mathfrak{S}(k+l)} \operatorname{sgn}\sigma \cdot \operatorname{sgn}\tau \cdot \xi(v^*_{\sigma\circ\tau(1)}, \cdots, v^*_{\sigma\circ\tau(k)})$$

$$\cdot \eta(v^*_{\sigma\circ\tau(k+1)}, \cdots, v^*_{\sigma\circ\tau(k+l)}) \cdot \varsigma(v^*_{\sigma\circ\tau(k+l+1)}, \cdots, v^*_{\sigma\circ\tau(k+l+h)})$$

$$= A_{k+l+h}(\xi \otimes \eta \otimes \varsigma)(v_1^*, \cdots, v_k^*, v^*_{k+1}, \cdots, v^*_{k+l}, v^*_{k+l+1}, \cdots, v^*_{k+l+h}),$$

故

$$(\xi ? \eta) ? \varsigma = A_{k+l+h}(\xi \otimes \eta \otimes \varsigma) = \xi ? (\eta ? \varsigma).$$

定理 6.2.6 外积有如下性质:

(1) 若 $\xi, \eta \in V = \Lambda^1(V)$,则

$$\xi ? \xi = 0, \quad \xi ? \eta = -\eta ? \xi.$$

一般地,若一个外积多项式含有两个以上相同的一次因子,则该式必为 0.

(2) 设 $\{e_1, \cdots, e_n\}$ 是 V 的一个基,则

$$e_{i_1} ? \cdots ? e_{i_r} = A_r(e_{i_1} \otimes \cdots \otimes e_{i_r}), \quad 1 \leqslant i_1, \cdots, i_r \leqslant n$$

是 $\Lambda^r(V)$ 的基,$r \leqslant n$. 注意到,上式只有当 $i_1, \cdots, i_r$ 互不相同时才不为 0. 当 $r > n$ 时,上式必为 0.

(3) $\Lambda^r(V) = \{0\}$, $\forall\, r > n$.

(4) $\forall\, r \leqslant n$,在 $\Lambda^r(V)$ 的基 $\{e_{j_1} ? \cdots ? e_{j_r}, 1 \leqslant j_1 < \cdots < j_r \leqslant n\}$ 之下,$\forall\, \xi \in \Lambda^r(V)$,都有表示式 $\xi = r! \sum\limits_{1 \leqslant i_1 < \cdots < i_r \leqslant n} \xi^{j_1, \cdots, j_r} e_{j_1} ? \cdots ? e_{j_r}$.

证 只要证明 $\{e_{j_1} ? \cdots ? e_{j_r}, 1 \leqslant j_1 < \cdots < j_r \leqslant n\}$ 中,共有 $\binom{n}{r} = \frac{n!}{r!(n-r)!}$ 个线性无关的外矢量. 分以下三步:①证 $e_{j_1} ? \cdots ? e_{j_r}$ 的求值公式. 任取 $v_1^*, \cdots, v_r^* \in V^*$,则

$$e_{j_1} ? \cdots ? e_{j_r}(v_1^*, \cdots, v_r^*) = \frac{1}{r!} \sum_{\sigma \in \mathfrak{S}(r)} \operatorname{sgn}\sigma \cdot \langle e_{j_1}, v^*_{\sigma(1)} \rangle \cdots \langle e_{j_r}, v^*_{\sigma(r)} \rangle$$

$$= \frac{1}{r!} \begin{vmatrix} \langle e_{j_1}, v_1^* \rangle & \cdots & \langle e_{j_1}, v_r^* \rangle \\ \langle e_{j_2}, v_1^* \rangle & \cdots & \langle e_{j_2}, v_r^* \rangle \\ \vdots & & \vdots \\ \langle e_{j_r}, v_1^* \rangle & \cdots & \langle e_{j_r}, v_r^* \rangle \end{vmatrix}. \tag{6.2.21}$$

上式就是 $e_{j_1} ? \cdots ? e_{j_r}$ 的求值公式,特别地,有

$$e_{j_1}?\cdots?e_{j_r}(e^*_{k_1},\cdots,e^*_{k_r})=\frac{1}{r!}\det[\langle e_{j_\alpha},e^*_{k_\beta}\rangle]=\frac{1}{r!}\delta^{j_1\cdots j_r}_{k_1\cdots k_r},\qquad(6.2.22)$$

其中

$$\delta^{j_1\cdots j_r}_{k_1\cdots k_r}=\begin{cases}1, & 1\leqslant j_1<\cdots<j_r\leqslant n,\{k_1,\cdots,k_r\}\text{ 是}\{j_1,\cdots,j_r\}\text{ 的偶置换},\\ -1, & 1\leqslant j_1<\cdots<j_r\leqslant n,\{k_1,\cdots,k_r\}\text{ 是}\{j_1,\cdots,j_r\}\text{ 的奇置换},\\ 0, & \text{其他}\end{cases}$$

称为广义 Kronecker 符号.

② 由①得 $e_1?\cdots?e_n\neq0$.

③ 当 $r\leqslant n$ 时,若$\{e_{j_1}?\cdots?e_{j_r},1\leqslant j_1<\cdots<j_r\leqslant n\}$线性相关,则 $\exists\,\alpha^{j_1,\cdots,j_r}\in\mathbb{F}$ 不全为 0,使 $\sum\limits_{1\leqslant j_1<\cdots<j_r\leqslant n}\alpha^{j_1,\cdots,j_r}e_{j_1}?\cdots?e_{j_r}=0$. 设其中一个不为 0 的常数为 $\overline{\alpha^{j_1,\cdots,j_r}}\neq0,1\leqslant j_1<\cdots<j_r\leqslant n$,加入与之相补的坐标指标为 $k_1,\cdots,k_{n-r}$,使得 $\{j_1,\cdots j_k,k_1,\cdots,k_{n-r}\}\in\mathfrak{T}(n)$,并用 $e_{k_1}?\cdots?e_{k_{n-r}}$ 外乘上式,得

$$\overline{\alpha^{j_1,\cdots,j_r}}\,e_{j_1}?\cdots?e_{j_r}?e_{k_1}?\cdots?e_{k_{n-k}}=\pm\overline{\alpha^{j_1,\cdots,j_r}}\,e_1?\cdots?e_n=0;$$

因 $e_1?\cdots?e_n\neq0$,只有$\overline{\alpha^{j_1,\cdots,j_r}}=0$,这与其不为 0 矛盾. 故$\{e_{j_1}?\cdots?e_{j_r},1\leqslant j_1<\cdots<j_r\leqslant n\}$线性无关. 于是得到$\{e_{j_1}?\cdots?e_{j_r},1\leqslant j_1<\cdots<j_r\leqslant n\}$是 $\Lambda^r(V)$的基,$\Lambda^r(V)$是$\begin{pmatrix}n\\r\end{pmatrix}=\dfrac{n!}{r!(n-r)!}$维线性空间.

至此,(1),(2),(4)都可由①,②,③得到.

设 r 次外矢量 ξ 的分量的坐标表示为 $\xi=\xi^{j_1,\cdots,j_r}e_{j_1}\otimes\cdots\otimes e_{j_r}$. 因为反对称化算子的线性性,故

$$\xi=A_r\xi=\xi^{j_1,\cdots,j_r}A_r(e_{j_1}\otimes\cdots\otimes e_{j_r})=\xi^{j_1,\cdots,j_r}e_{j_1}?\cdots?e_{j_r},$$

由(2)知,$\forall\,r>n\Rightarrow\Lambda^r(V)=\{0\}$. 此即(3). 定理得证.

2) 反对称 r 阶协变张量空间 $\Lambda^r(V^*)$中的外积运算

对于 $\Lambda^r(V^*)$,外积运算的定义也由定义 6.2.8 给出,并且可类似得到以下定理.

定理 6.2.7 外积有如下运算规则:对于 $\xi,\xi_1,\xi_2\in\Lambda^k(V^*),\eta,\eta_1,\eta_2\in\Lambda^l(V^*)$ 与 $\varsigma\in\Lambda^h(V^*)$,有

(1) **结合律** $(\xi?\eta)?\varsigma=\xi?(\eta?\varsigma)$;

(2) **分配律** $(\xi_1+\xi_2)?\eta=\xi_1?\eta+\xi_2?\eta,\qquad\xi?(\eta_1+\eta_2)=\xi?\eta_1+\xi?\eta_2$;

(3) **反交换律** $\xi?\eta=(-1)^{kl}\eta?\xi$.

定理 6.2.8 外积有如下性质:

(1) 若 $\xi,\eta\in V^*=\Lambda^1(V^*)$,则

$$\xi?\xi=0,\quad\xi?\eta=-\eta?\xi;$$

一般地,若一个外积多项式含有两个以上的相同的一次因子,则该式必为 0.

(2) 设$\{e_1^*,\cdots,e_n^*\}$是V^*的一个基,则

$$e_{i_1}^* ? \cdots ? e_{i_r}^* = A_r(e_{i_1}^* \otimes \cdots \otimes e_{i_r}^*), \quad 1 \leqslant i_1,\cdots,i_r \leqslant n$$

是$\Lambda^r(V^*)$的基,$r \leqslant n$. 注意到,上式只有$i_1,\cdots,i_r$互不相同时才不为 0. 当$r>n$时,上式必为 0.

(3) $\Lambda^r(V^*)=\{0\}, \forall r>n$.

(4) $\forall r \leqslant n$,在$\Lambda^r(V^*)$的基$\{e_{j_1}^* ? \cdots ? e_{j_r}^*, 1 \leqslant j_1 < \cdots < j_r \leqslant n\}$之下,对于$\forall \eta \in \Lambda^r(V^*)$都有表示$\eta = r! \sum\limits_{1 \leqslant k_1 < \cdots < k_r \leqslant n} \xi_{k_1,\cdots,k_r} e_{k_1}^* ? \cdots ? e_{k_r}^*$.

定理 6.2.8 的证明可仿照定理 6.2.6 的证明给出.

6.2.3 Grassmann 代数

设n维线性空间V及其对偶空间V^*的基分别为$\{e_1,\cdots,e_n\}$与$\{e_1^*,\cdots,e_n^*\}$.

1. 外矢量空间 $\Lambda(V)$,外形式空间 $\Lambda(V^*)$——Grassmann 代数(外代数)

定义 6.2.10(外矢量空间 $\Lambda(V)$) 记r次外矢量空间$\Lambda^r(V)(r=0,1,\cdots,n)$的分层和为

$$\Lambda(V) = \sum_{r=0}^{n} \Lambda^r(V), \tag{6.2.23}$$

则$\Lambda(V)$是2^n维线性空间. 对于$\xi,\eta \in \Lambda(V)$,有$\xi = \sum\limits_{r=0}^{n} \xi_r, \xi_r \in \Lambda^r(V)$与$\eta = \sum\limits_{s=0}^{n} \eta_s, \eta_s \in \Lambda^s(V)$. 定义外积$\xi ? \eta = \sum\limits_{r,s=0}^{n} \xi_r ? \eta_s$,则$(\Lambda(V),+,\alpha\cdot,?)$成为一个**外代数(Grassmann 代数)**,称为线性空间V上的**外矢量空间(outer vector space)**,其中的元称为**外矢量(outer vector)**.

$\Lambda(V)$的基为

$$\left\{\begin{matrix} 1;\ e_j : 1 \leqslant j \leqslant n;\ e_{j_1} ? e_{j_2} : 1 \leqslant j_1 < j_2 \leqslant n;\ \cdots; \\ e_{j_1} ? \cdots ? e_{j_r} : 1 \leqslant j_1 < \cdots < j_r \leqslant n;\ \cdots;\ e_1 ? \cdots ? e_n \end{matrix}\right\}. \tag{6.2.24}$$

外形式空间?(?*)可类似定义.

定义 6.2.11(外形式空间 $\Lambda(V^*)$) 对偶空间V^*上的外代数$\Lambda(V^*) = \sum\limits_{r=0}^{n} \Lambda^r(V^*)$,称为线性空间$V$上的**外形式空间(outer form space)**,其中的元称为**外形式(outer form)**.

$\Lambda(V^*)$的基为

$$\left\{\begin{matrix} 1;\ e_j^* : 1 \leqslant j \leqslant n;\ e_{j_1}^* ? e_{j_2}^* : 1 \leqslant j_1 < j_2 \leqslant n;\ \cdots; \\ e_{j_1}^* ? \cdots ? e_{j_r}^* : 1 \leqslant j_1 < \cdots < j_r \leqslant n;\ \cdots;\ e_1^* ? \cdots ? e_n^* \end{matrix}\right\}. \tag{6.2.25}$$

2. 空间 $\Lambda^r(V)$与空间 $\Lambda^r(V^*)$中元的配合

定义 6.2.12（$\Lambda^r(V)$与 $\Lambda^r(V^*)$中元的配合） 空间 $\Lambda^r(V)$与 $\Lambda^r(V^*)$是彼此对偶的，对于 $v_1?\cdots?v_r\in\Lambda^r(V)$与 $v_1^*?\cdots?v_r^*\in\Lambda^r(V^*)$，定义配合

$$\langle v_1?\cdots?v_r, v_1^*?\cdots?v_r^*\rangle = \det[\langle v_\alpha, v_\beta^*\rangle]. \tag{6.2.26}$$

注意到，$\Lambda^r(V)$的基为$\{e_{j_1}?\cdots?e_{j_r}:1\leqslant j_1<\cdots<j_r\leqslant n\}$，$\Lambda^r(V^*)$的基为$\{e_{j_1}^*?\cdots?e_{j_r}^*:1\leqslant j_1<\cdots<j_r\leqslant n\}$，则关于基的配合为

$$\begin{aligned}&\langle e_{j_1}?\cdots?v_{j_r}, e_{k_1}^*?\cdots?e_{k_r}^*\rangle\\&\quad=\det[\langle e_{j_\alpha}, e_{j_\beta}^*\rangle]\\&\quad=\delta_{j_1;\cdots;j_r}^{k_1;\cdots;k_r}=\begin{cases}1, & \{k_1,\cdots,k_r\}=\{j_1,\cdots,j_r\},\\0, & \{k_1,\cdots,k_r\}\neq\{j_1,\cdots,j_r\}.\end{cases}\end{aligned}$$

因此这两组基恰巧是彼此对偶的.

3. 两个光滑流形 V 与 W 的外形式空间之间的映射

回顾 6.1.1 节，对于 $f:C_p^\infty(V)\to C_q^\infty(W)$，相应地可类似定义 $T_p(V)$，$T_p(V^*)$；$T_p^*(V)$，$T_p^*(V^*)$；$T_q(W)$，$T_q(W^*)$；$T_p^*(W)$，$T_q^*(W^*)$，并且有以下关系：

光滑流形 V 与 W 之间的光滑映射 $F:V\to W$

↓

$F:V\to W$ 的微分 —— $F^*:T_q^*(W)\to T_p^*(V)$

$F^*((\mathrm{d}f))=\mathrm{d}(f\circ F)$，　$(\mathrm{d}f)\in T_q^*(W)$

↓

$F:V\to W$ 的切映射 —— $F_*:T_p(V)\to T_q(W)$

$\langle F_*X, a\rangle=\langle X, F^*a\rangle$，$X\in T_p(V)$，　$a\in T_q^*(W)$

现在对于两个 r 次**外形式空间** $\Lambda^r(W^*)$与 $\Lambda^r(V^*)$，可定义线性映射 $G:V\to W$ 的一个诱导映射 $G^*:\Lambda^r(W^*)\to\Lambda^r(V^*)$如下. 设 $v_1,\cdots,v_r\in V$ 与 $\varphi\in\Lambda^r(W^*)$，定义 G^* 为满足等式

$$G^*\circ\varphi(v_1,\cdots,v_r)\equiv\varphi(G(v_1),\cdots,G(v_r)) \tag{6.2.27}$$

的映射. 则 G^* 是线性的，且与外积运算?可交换. 于是，线性算子 G^* 是外代数$\Lambda(W^*)$到外代数$\Lambda(V^*)$的同态.

定理 6.2.9 设 $G:V\to W$ 是光滑线性映射，则由(6.2.27)式定义的线性算子 $G^*:\Lambda^r(W^*)\to\Lambda^r(V^*)$与外积运算?是可交换的，亦即，对于任意的 $\varphi\in\Lambda^r(W^*)$与 $\psi\in\Lambda^s(W^*)$，有

$$G^*(\varphi?\psi)=(G^*\varphi)?(G^*\psi).$$

4. 外积的几个重要性质

定理 6.2.10 矢量 $v_1,\cdots,v_r\in V$ 线性相关，当且仅当 $v_1?\cdots?v_r=0$.

证 必要性 若 $v_1,\cdots,v_r\in V$ 线性相关,可设 $v_r=\alpha_1 v_1+\cdots+\alpha_{r-1}v_{r-1}$,故

$$v_1 ? \cdots ? v_{r-1} ? (\alpha_1 v_1+\cdots+\alpha_{r-1}v_{r-1})$$

显然为 0.

充分性 反证,若 $v_1,\cdots,v_r$ 线性无关,则可将它扩充为 V 的基 $\{v_1,\cdots,v_r,u_{r+1},\cdots,u_n\}$,因为它们线性无关,从而 $v_1 ? \cdots ? v_r ? u_{r+1} ? \cdots ? u_n\neq 0$,所以 $v_1 ? \cdots ? v_r\neq 0$,这与 $v_1 ? \cdots ? v_r=0$ 矛盾. 定理得证.

定理 6.2.11(Cartan 引理) 设 $v_1,\cdots,v_r\in V$,$w_1,\cdots,w_r\in V$ 是 V 中的两组元,使得 $\sum\limits_{k=1}^{r} v_k ? w_k=0$,若 $v_1,\cdots,v_r$ 线性无关,则 w_k 可表示成线性组合 $w_k=\sum\limits_{j=1}^{r} a_{kj}v_j$,$1\leqslant k\leqslant r$,且 $a_{kj}=a_{jk}$.

证 因为 $v_1,\cdots,v_r$ 线性无关,可将其扩充成 V 的基 $\{v_1,\cdots,v_r,v_{r+1},\cdots,v_n\}$,可设

$$w_k=\sum_{j=1}^{r} a_{kj}v_j+\sum_{j=r+1}^{n} b_{kj}v_j,\quad k=1,2,\cdots,r,$$

代入 $\sum\limits_{k=1}^{r} v_k ? w_k=0$,得到

$$\begin{aligned}0&=\sum_{k,j=1}^{r} a_{kj}v_k ? v_j+\sum_{k=1}^{r}\sum_{i=r+1}^{n} b_{ki}v_k ? v_i\\&=\sum_{1\leqslant k<j\leqslant 1}^{r}(a_{kj}-a_{jk})v_k ? v_j+\sum_{k=1}^{r}\sum_{i=r+1}^{n} b_{ki}v_k ? v_i.\end{aligned}$$

由于 $\{v_k ? v_j,v_k ? u_j,u_k ? u_j:1\leqslant j<k\leqslant n\}$ 是 $\Lambda^2(V)$ 的一个基,因此上式给出

$$a_{kj}-a_{jk}=0 \quad 与 \quad b_{ki}=0,$$

此即 $w_k=\sum\limits_{j=1}^{r} a_{kj}v_j\,(k=1,2,\cdots,r)$ 与 $a_{kj}=a_{jk}$.

定理 6.2.12 设 $v_1,\cdots,v_r\in V$ 线性无关,w 是 V 上的 s 次外矢量,则存在 $\psi_1,\cdots,\psi_r\in\Lambda^{s-1}(V)$,使得 $w=\sum\limits_{k=1}^{r} v_k ? \psi_k$,当且仅当 $v_1 ? \cdots ? v_r ? w=0$.

定理 6.2.13 设 $v_j,w_j;v_j',w_j'\in V$,$1\leqslant j\leqslant k$,是空间 V 中的两组元,若 $\{v_j,w_j,1\leqslant j\leqslant k\}\subset V$ 是线性无关的,并且 $\sum\limits_{j=1}^{k} v_j ? w_j=\sum\limits_{j=1}^{k} v_j' ? w_j'$,则 $v_j',w_j'\in V$ 都是 $v_1,\cdots,v_k,w_1,\cdots,w_k$ 的线性组合,且它们也是线性无关的.

注 外积与行列式有密切关系. 由(6.2.21)式,有

$$e_{j_1} ? \cdots ? e_{j_r}(v_1^*,\cdots,v_r^*)=\frac{1}{r!}\sum_{\sigma\in\mathfrak{T}(r)}\operatorname{sgn}\sigma\cdot\langle e_{j_1},v_{\sigma(1)}^*\rangle\cdot\cdots\cdot\langle e_{j_r},v_{\sigma(r)}^*\rangle$$

$$=\frac{1}{r!}\begin{vmatrix}\langle e_{j_1},v_1^*\rangle & \cdots & \langle e_{j_1},v_r^*\rangle \\ \langle e_{j_2},v_1^*\rangle & \cdots & \langle e_{j_2},v_r^*\rangle \\ \vdots & & \vdots \\ \langle e_{j_r},v_1^*\rangle & \cdots & \langle e_{j_r},v_r^*\rangle\end{vmatrix}.$$

进而，设 $v_1,\cdots,v_k\in V$，且 $w_1,\cdots,w_k\in V$ 是前者的线性组合，亦即 $w_j=\sum_{\beta=1}^{k}t_j^{\beta}v_{\beta}$，则

$$w_1?\cdots?w_k=\det[t_j^{\beta}]v_1?\cdots?v_k, \tag{6.2.28}$$

因此，外矢量 $w_1?\cdots?w_k$ 与 $v_1?\cdots?v_k$ 只相差一个行列式作为数量因子.

6.3　外微分

6.3.1　张量丛、矢量丛

1. (r,s)型张量空间 $T_s^r(p)$

6.2 节中定义了(r,s)型张量空间，$V_s^r=\underbrace{V\otimes\cdots\otimes V}_{r}\otimes\underbrace{V^*\otimes\cdots\otimes V^*}_{s}$. 本节将张量理论用于光滑流形.

给定 m 维光滑流形 $M\equiv(M,\mathbb{A})$，取 $p\in M$，将 V,V^* 分别取作点 p 的切空间 T_p 与余切空间 T_p^*. 于是，在流形 M 的每一点 $p\in M$，有一个(r,s)型张量空间 $T_s^r(p)=\underbrace{T_p\otimes\cdots\otimes T_p}_{r}\otimes\underbrace{T_p^*\otimes\cdots\otimes T_p^*}_{s}$，它是一个 m^{r+s}维线性空间. 由此将构造几个重要的空间.

V 为实数域$\mathbb{R}$上的 m 维线性空间，$\{e_1,\cdots,e_m\}$为 V 的一个基；于是

$$y=(y_1,\cdots,y_m)\in V\leftrightarrow\mathbb{R}^m;$$

$GL(V)$为 V 上的线性自同构群；于是

$$GL(V)\leftrightarrow GL(m;\mathbb{R})\leftrightarrow\mathfrak{M}_{m\times m}=\{A=[a_{jk}]_{m\times m}:\det[a_{jk}]\neq 0\},$$

且 $A\in GL(m;\mathbb{R})$对 $y\in V$ 的作用记为

$$y\cdot A=(y_1,\cdots,y_m)\cdot[a_{jk}],\quad \det[a_{jk}]\neq 0,$$

这里 · 是矩阵的乘法；

V_s^r 为 V 上的(r,s)型张量空间，$V_s^r=\underbrace{V\otimes\cdots\otimes V}_{r}\otimes\underbrace{V^*\otimes\cdots\otimes V^*}_{s}$，$V_s^r$ 的基为

$$e_{j_1}\otimes\cdots\otimes e_{j_r}\otimes e_{k_1}^*\otimes\cdots\otimes e_{k_s}^*,\quad 1\leqslant j_\alpha,k_\beta\leqslant m,1\leqslant\alpha\leqslant r,\ 1\leqslant\beta\leqslant s;$$

T_p^*为具有局部坐标系$\{(U,u_j)\}$的、m 维光滑流形 M 在点 $p\in M$ 的余切空间，其自然基与对偶基(即切空间 T_p 的自然基)分别为

$$\{(\mathrm{d}u_k)_p,1\leqslant k\leqslant m\}\subset T_p^*\quad 与\quad \left\{\left.\frac{\partial}{\partial u_j}\right|_p,1\leqslant j\leqslant m\right\}\subset T_p;$$

$T_s^r(p)$为点 p 的(r,s)型张量空间,$T_s^r(p)=\underbrace{T_p\otimes\cdots\otimes T_p}_{r}\otimes\underbrace{T_p^*\otimes\cdots\otimes T_p^*}_{s}$,$T_s^r(p)$的基为

$$\left\{\begin{array}{c}\left(\dfrac{\partial}{\partial u_{j_1}}\right)_p\otimes\cdots\otimes\left(\dfrac{\partial}{\partial u_{j_r}}\right)_p\otimes(\mathrm{d}u_{k_1})_p\otimes\cdots\otimes(\mathrm{d}u_{k_s})_p,\\ 1\leqslant j_\alpha,k_\beta\leqslant m,\ 1\leqslant\alpha\leqslant r,\ 1\leqslant\beta\leqslant s.\end{array}\right\}$$

2. (r,s)型张量丛 $T_s^r\equiv T_s^r(M)$

令

$$T_s^r\equiv T_s^r(M)=\bigcup_{p\in M}T_s^r(p),$$

其中 $T_s^r(p)=\underbrace{T_p\otimes\cdots\otimes T_p}_{r}\otimes\underbrace{T_p^*\otimes\cdots\otimes T_p^*}_{s}$为$(r,s)$型张量空间,这里“$\bigcup\limits_{p\in M}$”称为隔离并. 我们要在 T_s^r 上引进拓扑结构,使 T_s^r 成为有可数基的 Hausdorff 空间,也要引进 T_s^r 上的C^∞微分结构,使其成为一个光滑流形,并且证明 T_s^r 微分同胚于一个积流形 $M\times\mathbb{R}^\mu$. 把 T_s^r 称为流形 M 上的(r,s)型张量丛.

1) T_s^r上的拓扑结构 对于实数域$\mathbb{R}$上的 m 维光滑流形 M,设$\{(U,u_j)\}$为一个局部坐标系,取 m 维线性空间 V(实际上就可视 V 为$\mathbb{R}^m$),$\{e_1,\cdots,e_m\}$为 V 的一组基.

定义映射

$$\varphi_U:U\times V_s^r\to\bigcup_{p\in U}T_s^r(p),\tag{6.3.1}$$

满足$\forall(p,y)\in U\times V_s^r\Rightarrow\varphi_U(p,y)\in T_s^r(p)$,使得$\varphi_U(p,y)$关于 $T_s^r(p)$的基

$$\left\{\left(\frac{\partial}{\partial u_{j_1}}\right)_p\otimes\cdots\otimes\left(\frac{\partial}{\partial u_{j_r}}\right)_p\otimes(\mathrm{d}u_{k_1})_p\otimes\cdots\otimes(\mathrm{d}u_{k_s})_p,\quad 1\leqslant j_\alpha,k_\beta\leqslant m\right\}$$

的表示

$$\varphi_U(p,y)=z_{k_1,\cdots,k_s}^{j_1,\cdots,j_r}\left(\left(\frac{\partial}{\partial u_{j_1}}\right)_p\otimes\cdots\otimes\left(\frac{\partial}{\partial u_{j_r}}\right)_p\otimes(\mathrm{d}u_{k_1})_p\otimes\cdots\otimes(\mathrm{d}u_{k_s})_p\right)$$

与 $y\in V_s^r$ 关于基

$$e_{j_1}\otimes\cdots\otimes e_{j_r}\otimes e_{k_1}^*\otimes\cdots\otimes e_{k_s}^*,\quad 1\leqslant j_\alpha,k_\beta\leqslant m$$

的表示

$$y=y_{k_1,\cdots,k_s}^{j_1,\cdots,j_r}(e_{j_1}\otimes\cdots\otimes e_{j_r}\otimes e_{k_1}^*\otimes\cdots\otimes e_{k_s}^*)$$

相同. 显然,$(p,y)\xleftrightarrow{\text{一对一}}\varphi_U(p,y)$.

取 M 的一个开覆盖$\{U,W,\cdots\}$,并且相应于每一个局部坐标,都定义(6.3.1)式中的映射,于是得到

$$\{\varphi_U,\varphi_W,\cdots\}.\tag{6.3.2}$$

把诸如 $U\times V_s^r$ 的各开子集在 φ_U 下的象的全体所成的集合作为 T_s^r 的拓扑基(不难验证,所得到的集合满足拓扑基的条件). 于是,这样确定了 T_s^r 的拓扑,使得 T_s^r 成为具有可数基的

Hausdorff 空间，并且每一个由(6.3.1)式定义的映射都是拓扑同胚.

2) T_s^r 上的 C^∞ 微分流形结构 对于任一固定点 $p\in U$，由(6.3.1)式中的 φ_U，定义一个映射 $\varphi_{U,p}:V_s^r\to T_s^r(p)$，满足

$$\varphi_{U,p}(y)=\varphi_U(p,y),\quad y\in V_s^r.$$

显然，$\varphi_{U,p}$是 V_s^r 到 $T_s^r(p)$ 的线性同构(即线性、双方单值、双方连续).

对于$\{U,W,\cdots\}$中的 $U\cap W\neq\varnothing$，且 $p\in U\cap W$，则由

$$V_s^r\xrightarrow[\varphi_{U,p}]{}T_s^r(p)\xrightarrow[\varphi_{W,p}^{-1}]{}V_s^r,$$

定义映射

$$g_{UW}(p)=\varphi_{W,p}^{-1}\circ\varphi_{U,p}:V_s^r\to V_s^r,\tag{6.3.3}$$

则 g_{UW}是 V_s^r 到 V_s^r 的自同构，从而 $g_{UW}(p)\in GL(V_s^r)$，并满足

$$y,y'\in V_s^r\Rightarrow\varphi_U(p,y)=\varphi_W(p,y');$$

而且可推得

$$\forall\, y,y'\in V_s^r\Rightarrow\varphi_U(p,y)=\varphi_W(p,y')\Leftrightarrow\varphi_{U,p}(y)=\varphi_{W,p}(y')$$
$$\Leftrightarrow y'=\varphi_{W,p}^{-1}\circ\varphi_{U,p}(y)\Leftrightarrow y'=y\cdot g_{UW}(p),$$

这最后一步是因为 $g_{UW}(p)\in GL(V_s^r)$. 这里 $y\equiv y_U\in U$，$y'\equiv y_W\in W$.

从另一个观点看映射 g_{UW}，得到

$$g_{UW}:U\cap W\to GL(V_s^r).\tag{6.3.4}$$

事实上，由 $\varphi_U:U\times V_s^r\to\bigcup\limits_{p\in U}T_s^r(p)$定义一个新映射 $\varphi_U(p,y)\equiv\varphi_{U,p}(y)$，使得 $\varphi_{U,p}$：$V_s^r\to T_s^r(p)$，因此(见图 6.3.1)

$$g_{UW}(p)=\varphi_{W,p}^{-1}\circ\varphi_{U,p}:V_s^r\to V_s^r.$$

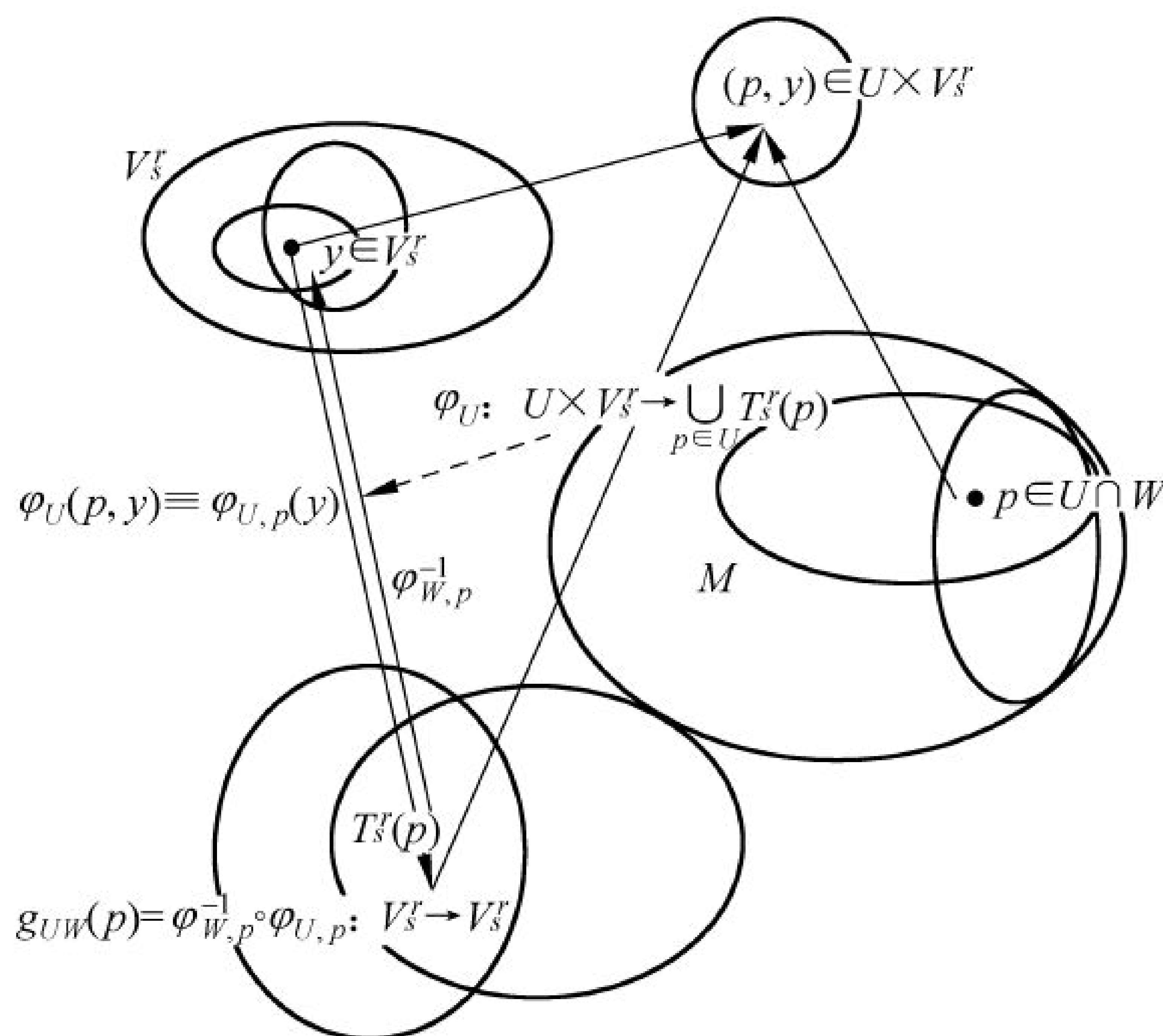


图 6.3.1 映射 $g_{UW}(p)$

亦即图 6.3.2 中的表示.

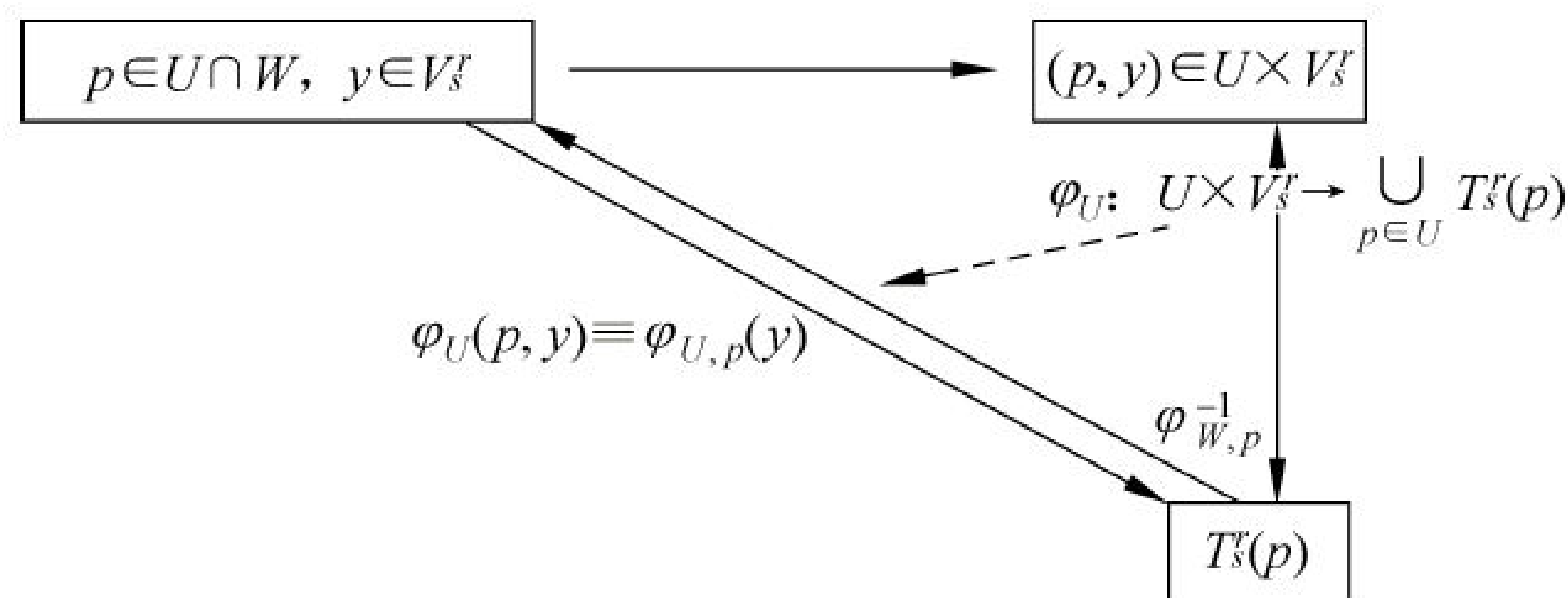


图 6.3.2 T_s^r 上的 C^∞ 微分流形结构

于是,有

$$p\in U\cap W\neq\varnothing\Rightarrow y,y'\in V_s^r,\quad (p,y),(p,y')\in U\times V_s^r$$
$$\Rightarrow\varphi_U(p,y)=\varphi_W(p,y')\Leftrightarrow y'=y\cdot g_{UW}(p).\tag{6.3.5}$$

我们需要证明以下两点:

(1) **g_{UW} 是光滑映射**

g_{UW} 的光滑性证明较为复杂,这里仅对 $r=1,s=1$ 情形加以证明.

设 M 关于局部坐标系 $\{(U,u_j)\}$ 的自然基为 $\{u_1,\cdots,u_m\}$,M 关于局部坐标系 $\{(W,w_j)\}$ 的自然基为 $\{w_1,\cdots,w_m\}$,切空间 T_p 与余切空间 T_p^* 的基分别为

$$\left\{\frac{\partial}{\partial u_j}\bigg|_p,1\leqslant j\leqslant m\right\}\subset T_p\quad 与\quad \{(\mathrm{d}u_k)_p,1\leqslant k\leqslant m\}\subset T_p^*,$$
$$\left\{\frac{\partial}{\partial w_j}\bigg|_p,1\leqslant j\leqslant m\right\}\subset T_p\quad 与\quad \{(\mathrm{d}w_k)_p,1\leqslant k\leqslant m\}\subset T_p^*.$$

设 $y,y'\in V_1^1$,用分量表示为 $y=y_k^j e_j\otimes e_k^*$, $y'=(y')_k^j e_j\otimes e_k^*$,因此

$$\begin{cases}\varphi_U(p,y)=y_k^j\left(\dfrac{\partial}{\partial u_j}\right)_p\otimes(\mathrm{d}u_k)_p,\\ \varphi_W(p,y')=(y')_k^j\left(\dfrac{\partial}{\partial w_j}\right)_p\otimes(\mathrm{d}w_k)_p.\end{cases}$$

于是,在 $U\cap W$ 上,自然基之间有以下关系:

$$\begin{cases}\mathrm{d}u_k=\dfrac{\partial u_k}{\partial w_k}\mathrm{d}w_k,\\ \dfrac{\partial}{\partial u_j}=\dfrac{\partial w_j}{\partial u_j}\dfrac{\partial}{\partial w_j},\end{cases}\quad 1\leqslant j,k\leqslant m;$$

而坐标变换的 Jacobi 矩阵为 $J_{UW}=\left[\left(\dfrac{\partial w_j}{\partial u_k}\right)_p\right]$,并且可以得到 $(y')_{k'}^{j'}=y_k^j\left[\left(\dfrac{\partial w_{j'}}{\partial u_j}\right)_p\right]\left[\left(\dfrac{\partial u_k}{\partial w_{k'}}\right)_p\right]$,亦即

$$(y\cdot g_{UW}(p))_{k'}^{j'}=y_k^j\left[\left(\frac{\partial w_{j'}}{\partial u_j}\right)_p\right]\left[\left(\frac{\partial u_k}{\partial w_{k'}}\right)_p\right].\tag{6.3.6}$$

事实上，$y\in V_1^1$ 的坐标应记为$(y_1^1,\cdots,y_1^m;y_2^1,\cdots,y_2^m;\cdots;y_m^1,\cdots,y_m^m)$，则(6.3.6)式表明，$g_{UW}(p)$用$(m^{r+s})^2=m^{2r}\times m^{2s}$阶非退化方阵表示，恰好是 Jacobi 方阵 J_{UW}与它的逆方阵J_{UW}^{-1}的张量积$J_{UW}\otimes J_{UW}^{-1}$，即

$$g_{UW}=J_{UW}\otimes J_{UW}^{-1}.$$

由于J_{UW}与$J_{UW}^{-1}=J_{WU}$在$U\cap W$上都是光滑的，因此g_{UW}也是光滑的.

(2) **T_s^r成为一个光滑流形**

由$\{\varphi_U(U\times V_s^r),\varphi_W(W\times V_s^r),\cdots\}$构成空间 T_s^r 的坐标开覆盖，并且是 C^∞ 相容的，故在每个坐标域 $\varphi_U(U\times V_s^r)$上的点(p,y)的坐标为$(u_i(p),y_{k_1,\cdots,k_s}^{j_1,\cdots,j_r})$，其中 u_i 是流形 M 的坐标域 U 上的局部坐标，$y_{k_1,\cdots,k_s}^{j_1,\cdots,j_r}$是 $y\in V_s^r$ 关于基 $e_{j_1}\otimes\cdots\otimes e_{j_r}\otimes e_{k_1}^*\otimes\cdots\otimes e_{k_s}^*$，$1\leqslant j_\alpha,k_\beta\leqslant m$ 的分量.

当 $U\cap W\neq\varnothing$时，由于 $g_{UW}=J_{UW}\otimes J_{UW}^{-1}$，故

$$g_{UW}=J_{UW}\otimes J_{UW}^{-1}:U\cap W\to GL(V_s^r)$$

是光滑的，关系式(6.3.5)说明，T_s^r 由$\{\varphi_U(U\times V_s^r),\varphi_W(W\times V_s^r),\cdots\}$构成的坐标开覆盖是 C^∞ 相容的，因此，T_s^r 成为一个光滑流形.

于是，其上的 C^∞ 微分流形结构为

$$\mathbb{A}=\{(U,\varphi_U)\}\equiv\{\varphi_U(U\times V_s^r),\ h\circ g_{UW}\},$$

其中 g_{UW}由(6.3.3)式或(6.3.4)式确定，$h:GL(V_s^r)\to\mathbb{R}^\mu$是 $GL(V_s^r)$到$\mathbb{R}^\mu$ 的拓扑同构，$\dim GL(V_s^r)=(m^{r+s})^2$.

定义 6.3.1(**(r,s)型张量丛、丛投影、纤维**) 隔离并 $T_s^r\equiv T_s^r(M)=\bigcup\limits_{p\in M}T_s^r(p)$在 C^∞ 微分流形结构 $\mathbb{A}=\{(U,\varphi_U)\}\equiv\{\varphi_U(U\times V_s^r),\ h\circ g_{UW}\}$之下成为一个光滑流形，称为**流形 M 上的(r,s)型张量丛(tensor bundle)**.

T_s^r 上的自然投影 $\pi:T_s^r\to M$ 把点 p 的(r,s)型张量空间

$$T_s^r(p)=\underbrace{T_p\otimes\cdots\otimes T_p}_{r}\otimes\underbrace{T_p^*\otimes\cdots\otimes T_p^*}_{s}$$

中的元映射到点 $p\in M$，则 $\pi:T_s^r\to M$ 是光滑满映射，称为**丛投影**，也称为**自然投影**；$T_s^r(p)$称为(r,s)**张量丛T_s^r在点 $p\in M$ 上的纤维(fibre)**.

将光滑流形 M 及其上的张量丛看成一体，恰似人的头盖及发丛，而纤维就相当一根头发丝，如图 6.3.3 所示.

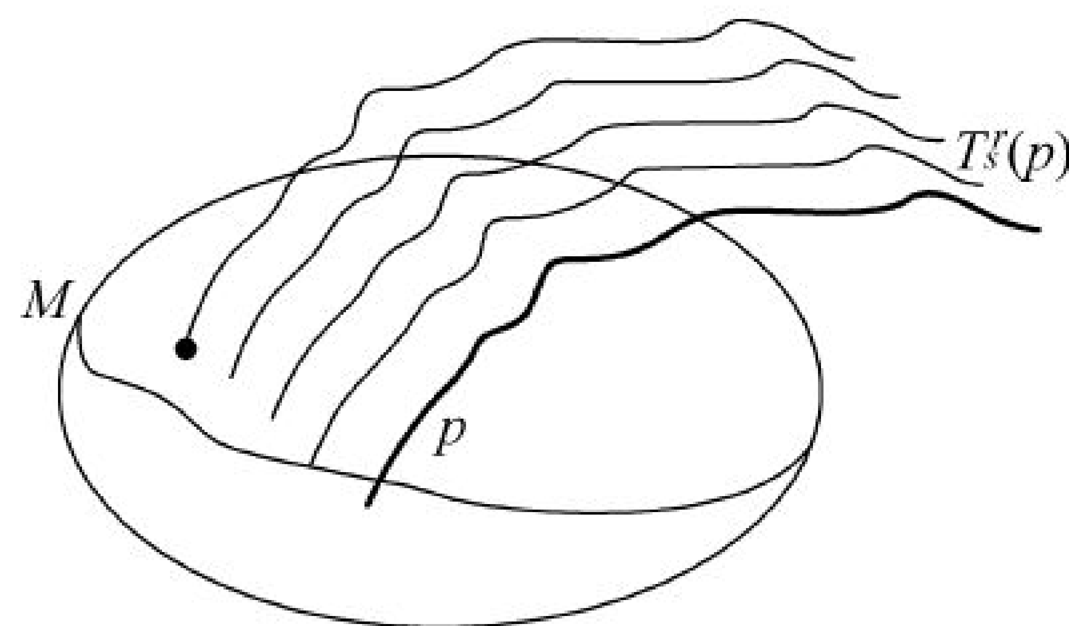


图 6.3.3 张量丛

当考虑两个光滑流形 M,N 之间的光滑映射时，就可定义一般的“(r,s)矢量丛”，而(r,s)张量丛是(r,s)矢量丛的特例. 张量丛、矢量丛在定义流形的“连络”时，成为“规范场理论”的重要数学基础，这里就不介绍了.

3. 切丛、余切丛

定义 6.3.2(切丛、余切丛) 在(r,s)张量丛 $T^r_s \equiv T^r_s(M) = \bigcup\limits_{p\in M} T^r_s(p)$ 中,令 $r=1,s=0$,就得到流形 M 上的$(1,0)$**型张量丛——切丛(tangent bundle)**,亦即,切丛表示为

$$T(M) \equiv T^1_0(M) = \bigcup_{p\in M} T^1_0(p) = \bigcup_{p\in M} T_p;$$

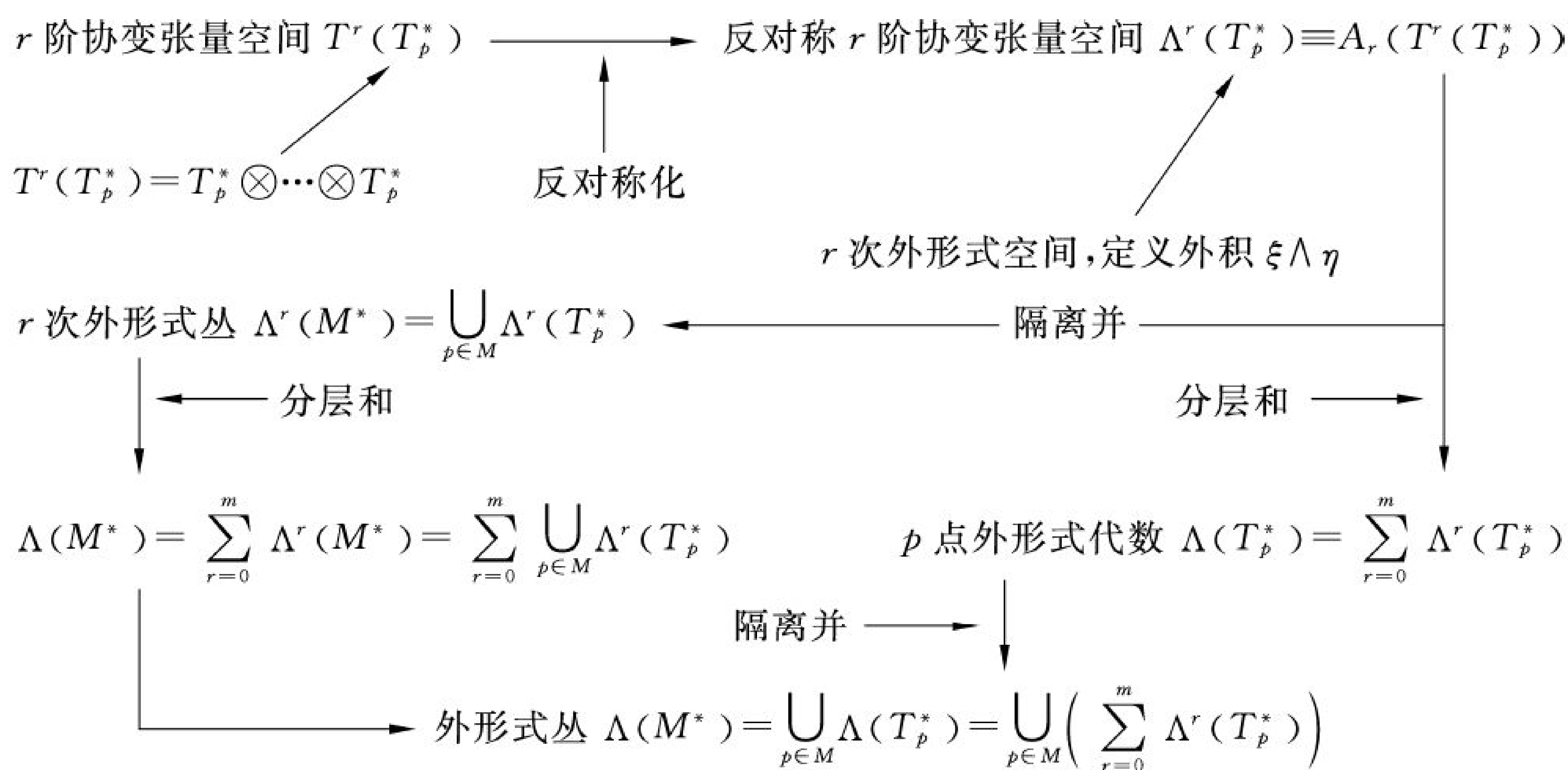
若令 $r=0,s=1$,就得到流形 M 上的$(0,1)$**型张量丛——余切丛(cotangent bundle)**,亦即,余切丛表示为

$$T^*(M) \equiv T^0_1(M) = \bigcup_{p\in M} T^0_1(p) = \bigcup_{p\in M} T^*_p.$$

4. 外矢量丛,外形式丛

取 $V=T_p, V^*=T^*_p$,按照张量丛的做法,可以在光滑流形 M 上构造外矢量丛与外形式丛. 我们以构造外形式丛为例.

定义 6.3.3(外形式丛) 建立外形式丛 $\Lambda(M^*)$的框图如下.



从以上构建"丛"与"代数"的过程看到,"丛"是在流形上构建"隔离并","代数"是在给定点处构建"分层和".

外矢量丛 $\Lambda(M)$的构造与以上外形式丛 $\Lambda(M^*)$的构造步骤类似,请读者自行完成.

5. 光滑截面

1) 张量丛 T^r_s 的光滑截面——(r,s)型光滑张量场

定义 6.3.4(张量丛 T^r_s 的光滑截面) 对张量丛 $T^r_s \equiv T^r_s(M) = \bigcup\limits_{p\in M} T^r_s(p)$,设 $g: M \to T^r_s$

是光滑映射，若 $g:M\to T_s^r$ 与丛投影 $\pi:T_s^r\to M$ 的复合 $\pi\circ g$ 是恒同映射，亦即 $\pi\circ g=I:M\to M$，则称 g 是**张量丛 T_s^r 的一个光滑截面(smooth section)**，或称 g 是 M **上的 (r,s) 型光滑张量场(tensor field)**.

图 6.3.4 中，虚线椭球形是流形 M，从流形 M 上出发的细曲线是**纤维** $T_s^r(p)$，粗曲线所围的是**光滑截面** f，图 6.3.4 整体就是一个张量丛 T_s^r.

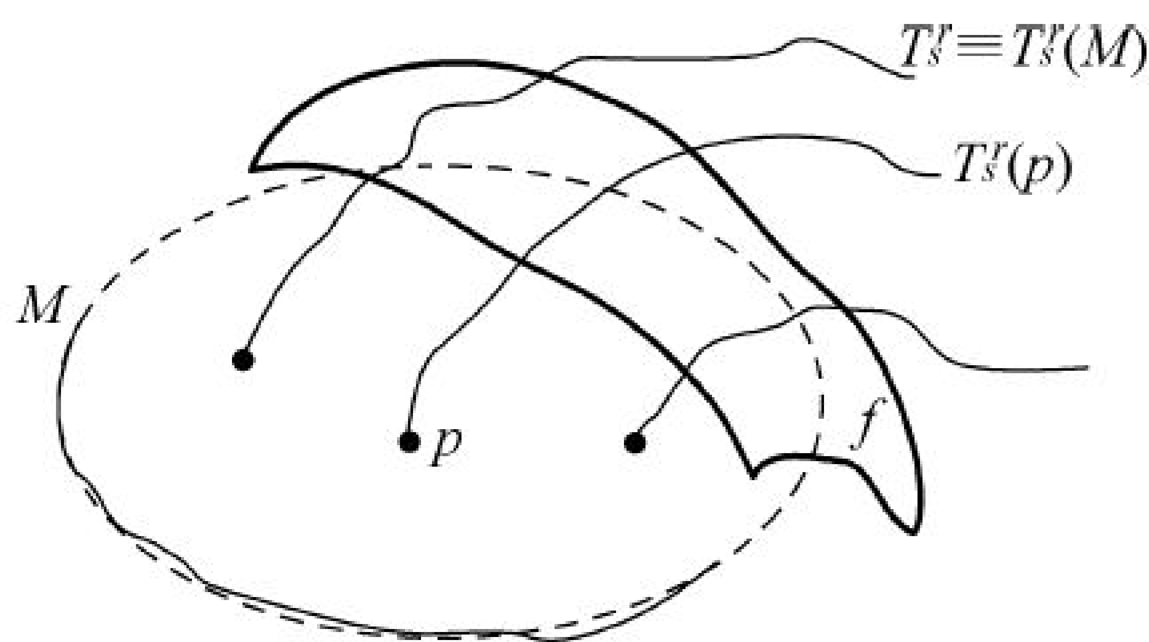


图 6.3.4　张量丛的光滑截面

实际上，g 在 $T_s^r(p)$ 上的像点就是 $g(p)$，像集 $g(M)$ 就是光滑截面，有

$$g(M)\subset T_s^r=\bigcup_{p\in M}T_s^r(p).$$

2）切丛 $T(M)$ 的光滑截面——切矢量场

定义 6.3.5(切丛 $T(M)$ 的光滑截面)　对于由在点 $p\in M$ 的切空间 T_p 构成的切丛 $T(M)\equiv T_0^1(M)=\bigcup_{p\in M}T_p$，设 $g:M\to T(M)$ 是光滑映射. 若 $g:M\to T(M)$ 与丛投影 $\pi:T(M)\to M$ 的复合 $\pi\circ g$ 是恒同映射，亦即 $\pi\circ g=I:M\to M$，则称 $g(M)$ 是**切丛 $T(M)$ 的一个光滑截面**，切丛的光滑截面就是 M 的**切矢量场**.

请读者回顾 6.1.1 节中定义的切矢量场，并进行比较.

定义 6.3.6(光滑切矢量场)　设 $X=\{X_p:p\in M\}$ 是光滑流形 M 上的一个切矢量场，若对于任意的 $f\in C^\infty(M)$，仍有 $Xf\in C^\infty(M)$，则称 X 是**流形 M 上的光滑切矢量场**.

于是，光滑切矢量场 $X=\{X_p:p\in M\}$ 是 $C^\infty(M)$ 到 $C^\infty(M)$ 的一个算子. 其性质如下：$\forall f,g\in C^\infty(M),\alpha,\beta\in\mathbb{R}$，有

(1) $X(\alpha f+\beta g)=\alpha(Xf)+\beta(Xg)$；

(2) $X(f\cdot g)=f\cdot(Xg)+g\cdot(Xf)$.

3）余切丛 $T^*(M)$ 的光滑截面——一次微分式

定义 6.3.7(余切丛 $T^*(M)$ 的光滑截面)　对于由在点 $p\in M$ 的余切空间 T_p^* 构成的余切丛 $T^*(M)\equiv T_1^0(M)=\bigcup_{p\in M}T_p^*$，设 $g:M\to T^*(M)$ 是光滑映射. 若映射 $g:M\to T^*(M)$ 与丛投影 $\pi:T^*(M)\to M$ 的复合 $\pi\circ g$ 是恒同映射，亦即 $\pi\circ g=I:M\to M$，则称 $g(M)$ 是**余切丛 $T^*(M)$ 的一个光滑截面**，余切丛的光滑截面就是 M 的**一次微分式**.

6.3.2　外微分式的外微分

我们从 m 维光滑流形 $M\equiv(M,\mathbb{A})$ 开始.

1. M 上的外微分式空间 $A(M)$

1) M 上的 r 次外微分式，r 次外微分式空间 $A^r(M)$

定义 6.3.3 给出了 M 上 r 次外形式丛 $\Lambda^r(M^*)=\bigcup_{p\in M}\Lambda^r(T_p^*)$，它是 M 上的一种张量丛. 记

$$A^r(M)=\Gamma(\Lambda^r(M^*))$$

为 $\Lambda^r(M^*)$的光滑截面的全体所成的空间，称 $A^r(M)$为 M 上的 **r 次外微分式空间**，其中的元称为 **M 上的 r 次外微分式(M 上的 r 次外微分式，就是光滑的反对称 r 阶协变张量场)**.

2) M 上的外微分式，外微分式空间 $A(M)$

定义 6.3.3 给出了 M 上的外形式丛 $\Lambda(M^*)=\bigcup_{p\in M}\Lambda(T_p^*)=\bigcup_{p\in M}\left(\sum_{r=0}^{m}\Lambda^r(T_p^*)\right)$，它是 M 上的一种张量丛. 记

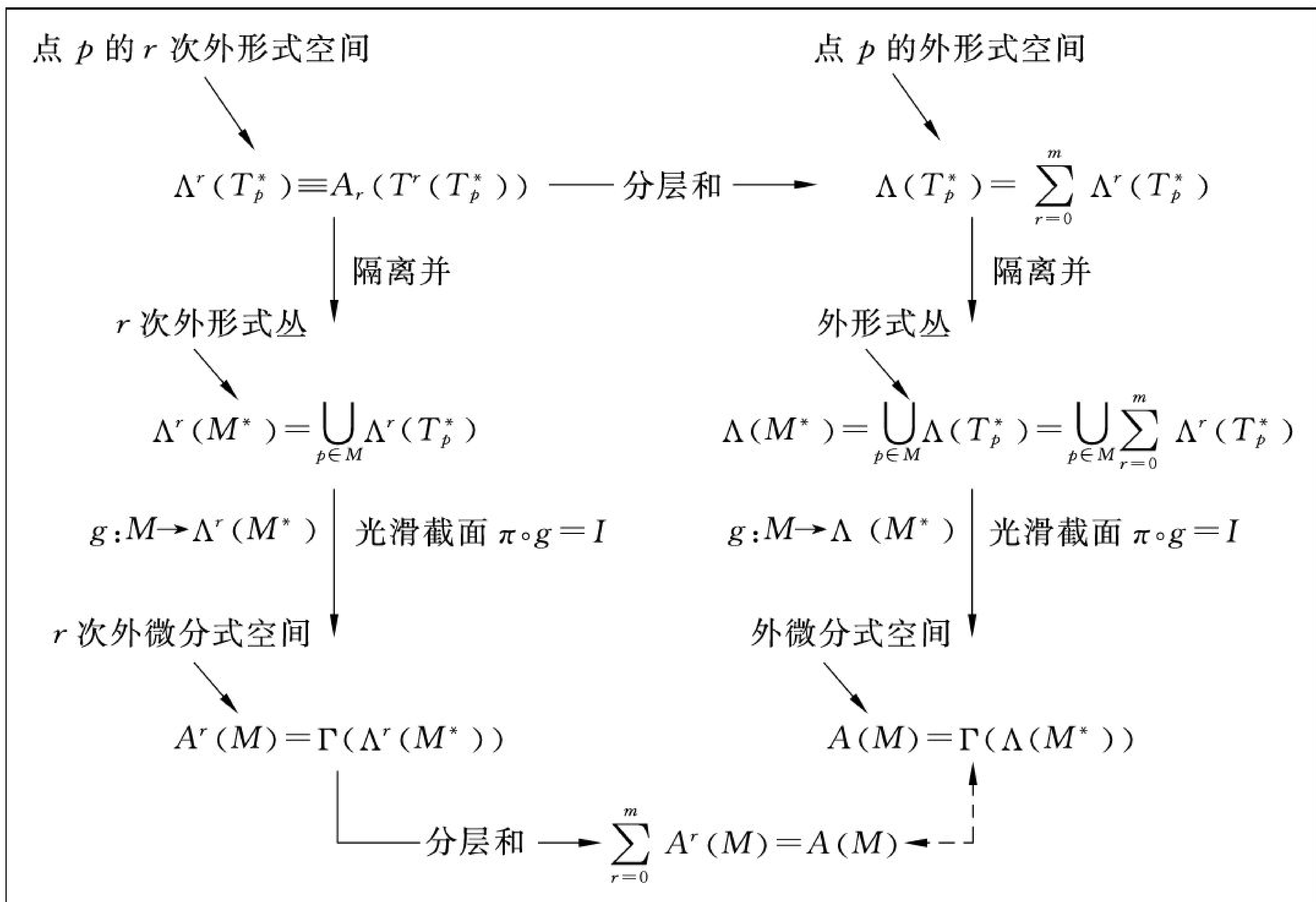
$$A(M)=\Gamma(\Lambda(M^*))$$

为 $\Lambda(M^*)$的光滑截面的全体所成的空间. 称 $A(M)$为 M 上的外微分式空间，其中的元称为 **M 上的外微分式(M 上的外微分式，就是光滑的反对称协变张量场)**.

注 1 $\Lambda^r(M^*)$是 r 次外微分式丛，亦即，未作分层和，故记号为 $A^r(M)=\Gamma(\Lambda^r(M^*))$.

注 2 $\Lambda(M^*)$是外微分式丛，已经作过分层和，没有"r 次"，故记号为 $A(M)=\Gamma(\Lambda(M^*))$.

3) 外微分式空间 $A(M)$的构成



从上面的交换图看到

$$A(M)=\sum_{r=0}^{m}A^r(M)=\sum_{r=0}^{m}\Gamma(\Lambda^r(M^*))=\sum_{r=0}^{m}\Gamma\left(\bigcup_{p\in M}\Lambda^r(T_p^*)\right),$$

也有

$$\begin{aligned}A(M)&=\Gamma(\Lambda(M^*))=\Gamma\left(\bigcup_{p\in M}\Lambda(T_p^*)\right)=\Gamma\left(\bigcup_{p\in M}\left(\sum_{r=0}^{m}\Lambda^r(T_p^*)\right)\right)\\&=\Gamma\left(\sum_{r=0}^{m}\left(\bigcup_{p\in M}\Lambda^r(T_p^*)\right)\right)=\sum_{r=0}^{m}\Gamma\left(\bigcup_{p\in M}\Lambda^r(T_p^*)\right).\end{aligned}$$

下面要用到两个空间的基.

(1) **点 p 的 r 次外形式空间** $\Lambda^r(T_p^*)\equiv A_r(T^r(T_p^*))$的基为

$$(\mathrm{d}u_{k_1})_p ?\cdots?(\mathrm{d}u_{k_r})_p,\quad 1\leqslant k_1,\cdots,k_r\leqslant m,$$

对 r 作分层直和 $\sum_{r=0}^{m}\Lambda^r(T_p^*)$,得到外代数$\left(\sum_{r=0}^{m}\Lambda^r(T_p^*),+,\alpha\cdot,?\right)$;

(2) 点 p 的 r **次外矢量空间** $\Lambda^r(T_p)\equiv A_r(T^r(T_p))$的基为

$$\left(\frac{\partial}{\partial u_{j_1}}\right)_p ?\cdots?\left(\frac{\partial}{\partial u_{j_r}}\right)_p,\quad 1\leqslant j_1,\cdots,j_r\leqslant m$$

对 r 作分层直和 $\sum_{r=0}^{m}\Lambda^r(T_p)$,得到外代数$\left(\sum_{r=0}^{m}\Lambda^r(T_p),+,\alpha\cdot,?\right)$.

4) 外微分式空间 $A(M)$ 中外微分式 $\omega\in A(M)=\sum_{r=0}^{m}A^r(M)$ 的表示

每个 $\omega\in A(M)=\sum_{r=0}^{m}A^r(M)$,可惟一地表示为

$$\omega=\omega^0+\omega^1+\cdots+\omega^m=\sum_{r=0}^{m}\omega^r,$$

其中 $\omega^r\in A^r(M)$为 r 次外微分式,$r\in\{0,1,\cdots,m\}$.

5) 外微分式空间 $A(M)$中的外积运算

现在将外积运算?推广到外微分式空间

$$A(M)=\sum_{r=0}^{m}A^r(M)=\sum_{r=0}^{m}\Gamma(\Lambda^r(M^*))=\sum_{r=0}^{m}\Gamma\left(\bigcup_{p\in M}\Lambda^r(T_p^*)\right).$$

我们有

$$\omega\in A(M)=\sum_{r=0}^{m}A^r(M)\Rightarrow\omega\in A^j(M)=\Gamma\left(\bigcup_{p\in M}\Lambda^j(T_p^*)\right),\quad 0\leqslant j\leqslant m,$$

亦即,ω 是点 $p\in M$ 的反对称 j 次协变张量空间 $\Lambda^j(T_p^*)(0\leqslant j\leqslant m)$形成的丛

$$\bigcup_{p\in M}\Lambda^j(T_p^*),\quad 0\leqslant j\leqslant m$$

的光滑截面 $\Gamma\left(\bigcup_{p\in M}\Lambda^j(T_p^*)\right)$,$0\leqslant j\leqslant m$. 因此 $\omega:M\to\bigcup_{p\in M}\Lambda^j(T_p^*)$,$0\leqslant j\leqslant m$,所以 $\omega(p)\in$

$\Lambda^j(T_p^*)$是一个反对称 j 次协变张量，亦即 j 次外形式.

于是，$\omega_1,\omega_2\in A(M)=\sum\limits_{r=0}^{m}A^r(M)$表示对任意点 $p\in M$，都有

$$\omega_1(p)\in\Lambda^j(T_p^*),\quad \omega_2(p)\in\Lambda^k(T_p^*),\quad 0\leqslant j,k\leqslant m,$$

故 $\omega_1(p)?\omega_2(p)$作为两个外形式 $\omega_1(p),\omega_2(p)$的外积有意义. 因此，可定义 $\omega_1?\omega_2$ 为 $(\omega_1?\omega_2)(p)\equiv\omega_1(p)?\omega_2(p)$，并且 $\omega_1?\omega_2\in A(M)$. 于是

$$\left(A(M)=\sum_{r=0}^{m}A^r(M),+,\alpha\cdot,\tau,?\right)$$

成为一个**分层代数(graded algebra)**，亦即，$A(M)$是一系列线性空间 $A^r(M)$的分层和 $A(M)=\sum\limits_{r=0}^{m}A^r(M)$. 外积?作为空间上的运算，给出

$$?:A^r(M)\times A^s(M)\to A^{r+s}(M)\equiv\begin{cases}A^{r+s}(M), & r+s\leqslant m,\\ \{0\}, & r+s>m.\end{cases}$$

6) r 次外矢量丛 $\Lambda^r(M)$与 r 次外形式丛 $\Lambda^r(M^*)$的配合

r 次外矢量丛 $\Lambda^r(M)=\bigcup\limits_{p\in M}\Lambda^r(T_p)$与 r 次外形式丛 $\Lambda^r(M^*)=\bigcup\limits_{p\in M}\Lambda^r(T_p^*)$是对偶丛，于是，在流形 M 的同一点 $p\in M$ 上的纤维之间的配合，是从 $\Lambda^r(T_p)$与 $\Lambda^r(T_p^*)$的配合诱导来的. 因此，空间 $\Lambda^r(T_p)$的基 $\left(\frac{\partial}{\partial u_{j_1}}\right)_p?\cdots?\left(\frac{\partial}{\partial u_{j_r}}\right)_p$，$1\leqslant j_1,\cdots,j_r\leqslant m$，与空间 $\Lambda^r(T_p^*)$的基 $(\mathrm{d}u_{k_1})_p?\cdots?(\mathrm{d}u_{k_r})_p$，$1\leqslant k_1,\cdots,k_r\leqslant m$，应满足配合关系

$$\left\langle\left(\frac{\partial}{\partial u_{j_1}}\right)_p?\cdots?\left(\frac{\partial}{\partial u_{j_r}}\right)_p,(\mathrm{d}u_{k_1})_p?\cdots?(\mathrm{d}u_{k_r})_p\right\rangle=\delta_{j_1\cdots j_r}^{k_1\cdots k_r},$$

而且 $\omega(p)\in\Lambda^r(T_p^*)$在局部坐标系$\{(U,u_j)\}$下的表示为 $\omega_{k_1\cdots k_r}(\mathrm{d}u_{k_1}?\cdots?\mathrm{d}u_{k_r})$，分量 $\omega_{k_1\cdots k_r}$ 可表示成

$$\omega_{k_1\cdots k_r}=\frac{1}{r!}\left\langle\left(\frac{\partial}{\partial u_{k_1}}\right)_p?\cdots?\left(\frac{\partial}{\partial u_{k_r}}\right)_p,\omega\right\rangle.$$

2. 外微分式空间 $A(M)$上的外微分运算及其性质

本小节中，将在外微分式空间$(A(M),+,\alpha\cdot,\mathbb{A},?)$上定义一种运算，称为外微分. 由于外微分的引入，使得 $A(M)$中又增加了一种运算，成为$(A(M),+,\alpha\cdot,\mathbb{A},?,\mathrm{d})$.

1) 外微分的定义

定义 6.3.8(外微分) 设 M 是 m 维光滑流形，$A(M)$是 M 上的外微分式空间，若存在映射 $\mathrm{d}:A(M)\to A(M)$，满足

(1) $\mathrm{d}(A^r(M))\subset A^{r+1}(M)$；

(2) 若 $\omega_1,\omega_2\in A(M)$，则 $\mathrm{d}(\omega_1+\omega_2)=\mathrm{d}\omega_1+\mathrm{d}\omega_2$；

(3) 若 ω_1 是 r 次外微分式，则 $\mathrm{d}(\omega_1 ? \omega_2)=(\mathrm{d}\omega_1) ? \omega_2+(-1)^r \omega_1 ? \mathrm{d}\omega_2$；

(4) 若 f 是 M 上的光滑函数，即 $f\in A^0(M)=C^\infty(M)$，则 $\mathrm{d}f$ 恰是 f 在经典意义下的微分；

(5) 若 $\omega\in A^0(M)$，则 $\mathrm{d}(\mathrm{d}\omega)=0$，

则称映射 $\mathrm{d}:A(M)\to A(M)$ 为外微分式空间 $A(M)$ 上的外微分，所确定的映射 d 称为**外微分算子**(**outward differential operator**)．对于 $\omega\in A(M)$，记 ω 的外微分为 $\mathrm{d}\omega$.

2) 外微分的存在性

下述定理给出外微分在外微分式空间 $(A(M),+,\alpha\cdot,\tau,?)$ 上的存在性.

定理 6.3.1(存在性定理)　设 M 是 m 维光滑流形，则存在惟一的映射

$$\mathrm{d}:A(M)\to A(M),$$

满足定义 6.3.8 中所述外微分的条件(1)～(5).

证　下面先定义 $\mathrm{d}:A(M)\to A(M)$，并证明它满足(1)～(5)．因此，外微分算子 d 存在.

① 首先指出，要定义的 d 显然是一个局部线性算子，亦即

$$\omega_1,\omega_2\in A(M),\ \omega_1|_U=\omega_2|_U,\ U\in\tau\Rightarrow \mathrm{d}\omega_1|_U=\mathrm{d}\omega_2|_U,$$

且

$$\omega\in A(M),\ \omega|_U=\omega|_W,\ U,W\in\tau\Rightarrow \mathrm{d}(\omega|_U)|_{U\cap W}=\mathrm{d}(\omega|_W)|_{U\cap W}.$$

② 其次证明所定义的“外微分算子 d”是惟一的. 具体证明过程如下.

设 $\omega\in A(M)$，则 $\exists\, r(0\leqslant r\leqslant m)$，$\omega\in A^r(M)\Rightarrow \omega|_U=a\cdot \mathrm{d}u_{k_1}?\cdots?\mathrm{d}u_{k_r}$，$a\in A^0(M)$ 是 $U\in\tau$ 上的光滑函数 $\Rightarrow$ 定义

$$\mathrm{d}\omega=\mathrm{d}(a\cdot \mathrm{d}u_{k_1}?\cdots?\mathrm{d}u_{k_r})\equiv \mathrm{d}a\ ?\mathrm{d}u_{k_1}?\cdots?\mathrm{d}u_{k_r},\tag{$*$}$$

其中

$$\mathrm{d}a=\frac{\partial a}{\partial u_j}\mathrm{d}u_j=\sum_{j=1}^{m}\frac{\partial a}{\partial u_j}\mathrm{d}u_j=\frac{\partial a}{\partial u_1}\mathrm{d}u_1+\cdots+\frac{\partial a}{\partial u_m}\mathrm{d}u_m,\tag{$**$}$$

且 $a=a(u_1,\cdots,u_m)\in C^\infty(M)$.

($*$)式与($**$)式 $\Leftrightarrow$ (4). 这是因为由 $a\in A^0(M)=C^\infty(M)$ 与($**$)式知 $\mathrm{d}a\in A^1(M)$，而 $A^1(M)$ 的基为 $\{\mathrm{d}u_1,\cdots,\mathrm{d}u_m\}$，因此有

$$\mathrm{d}a=\underset{\substack{\uparrow\\ \text{Einstein 记号}}}{\frac{\partial a}{\partial u_k}\mathrm{d}u_k}=\underset{\substack{\uparrow\\ \text{一次微分的表示}}}{\sum_{k=1}^{m}\frac{\partial a}{\partial u_j}\mathrm{d}u_j}$$

故(4)成立，反之亦然. 又由($*$)式与($**$)式知(2)显然成立；并且 $\forall\alpha\in\mathbb{R}$ 有 $\mathrm{d}(\alpha\omega)=\alpha\mathrm{d}\omega$，故所定义的 d 是线性算子. 下面，利用 $\omega\in A(M)$ 的局部坐标表示来验证(1)，(3)，(5).

验证(1)　由于 $\omega\in A^r(M)$，$\omega=a\cdot \mathrm{d}u_{k_1}?\cdots?\mathrm{d}u_{k_r}$，$a\in A^0(M)=C^\infty(M)$，根据($**$)式，$m$ 元光滑函数 a 的微分为 $\mathrm{d}a=\sum\limits_{k=1}^{m}\frac{\partial a}{\partial u_j}\mathrm{d}u_j$，因此，按($*$)式与($**$)式中的定义，对于 $\omega\in A(M)$，设 $\omega=a\cdot \mathrm{d}u_{k_1}?\cdots?\mathrm{d}u_{k_r}$，则有

$$\begin{aligned}
d\omega &= d(a \cdot du_{k_1} ? \cdots ? du_{k_r}) = da \,?\, du_{k_1} ? \cdots ? du_{k_r} \\
&= \left(\frac{\partial a}{\partial u_1} du_1 + \cdots + \frac{\partial a}{\partial u_m} du_m\right) ? du_{k_1} ? \cdots ? du_{k_r} \\
&= \frac{\partial a}{\partial u_1} du_1 ? (du_{k_1} ? \cdots ? du_{k_r}) + \cdots + \frac{\partial a}{\partial u_m} du_m ? (du_{k_1} ? \cdots ? du_{k_r}).
\end{aligned}$$

↑

$du_1 ? (du_{k_1} ? \cdots ? du_{k_r})$ 共 k_r+1 项，形成 $du_{k_1} ? \cdots ? du_{k_r+1}$ 项

于是，$\frac{\partial a}{\partial u_j} du_j ? (du_{k_1} ? \cdots ? du_{k_r}) \in A^{r+1}(M)$，$j \neq k_1, \cdots, k_r$，其余各项类似，故符合(1).

验证(3) 若 $\omega_1 = a \cdot du_{k_1} ? \cdots ? du_{k_r} \in A^r(M)$，$\omega_2 = b \cdot du_{j_1} ? \cdots ? du_{j_s} \in A^s(M)$，则按外积定义，有

$$\begin{aligned}
\omega_1 ? \omega_2 &= (a \cdot du_{k_1} ? \cdots ? du_{k_r}) ? (b \cdot du_{j_1} ? \cdots ? du_{j_s}) \\
&= (a \cdot b)(du_{k_1} ? \cdots ? du_{k_r} ? du_{j_1} ? \cdots ? du_{j_s}),
\end{aligned}$$

其中 $a \cdot b \in C^\infty(M)$，故按(*)式中的定义，有

$$\begin{aligned}
d(\omega_1 ? \omega_2) &= d\{(a \cdot du_{k_1} ? \cdots ? du_{k_r}) ? (b \cdot du_{j_1} ? \cdots ? du_{j_s})\} \\
&= d\{(a \cdot b)(du_{k_1} ? \cdots ? du_{k_r} ? du_{j_1} ? \cdots ? du_{j_s})\} \\
&= \{b(da) + a(db)\}(du_{k_1} ? \cdots ? du_{k_r} ? du_{j_1} ? \cdots ? du_{j_s}) \\
&= (b(da) ? du_{k_1} ? \cdots ? du_{k_r} ? du_{j_1} ? \cdots ? du_{j_s}) \\
&\quad + (a(db) ? du_{k_1} ? \cdots ? du_{k_r} ? du_{j_1} ? \cdots ? du_{j_s}) \\
&= ((da) ? du_{k_1} ? \cdots ? du_{k_r}) ? (b \,?\, du_{j_1} ? \cdots ? du_{j_s}) \\
&\quad + (-1)^r (a \,?\, du_{k_1} ? \cdots ? du_{k_r})((db) ? du_{j_1} ? \cdots ? du_{j_s}) \\
&= d\omega_1 ? \omega_2 + (-1)^r \omega_1 ? d\omega_2,
\end{aligned}$$

故符合(3).

验证(5) 若 $\omega = f \in A^0(M) = C^\infty(M)$，则 f 是 M 上的光滑函数，f 在 U 上的限制，由(**)式，有

$$df = \sum_{j=1}^{m} \frac{\partial f}{\partial u_j} du_j \Rightarrow$$

$$\begin{aligned}
d(df) &= d\left(\sum_{j=1}^{m} \frac{\partial f}{\partial u_j} du_j\right) ? (du_k) = \sum_{k=1}^{m} \sum_{j=1}^{m} \frac{\partial^2 f}{\partial u_j \partial u_k} (du_j ? du_k) \\
&= \sum_{k,j=1}^{m} \frac{1}{2} \left(\frac{\partial^2 f}{\partial u_j \partial u_k} + \frac{\partial^2 f}{\partial u_k \partial u_j}\right) du_j ? du_k \\
&= \frac{1}{2} \left(\sum_{j,k=1}^{m} \frac{\partial^2 f}{\partial u_j \partial u_k} du_j ? du_k - \sum_{k,j=1}^{m} \frac{\partial^2 f}{\partial u_k \partial u_j} du_k ? du_j\right) = 0,
\end{aligned}$$

最后一步是因为 $f\in C^{\infty}(M)$蕴含$\dfrac{\partial^2 f}{\partial u_j \partial u_k}=\dfrac{\partial^2 f}{\partial u_k \partial u_j}$,故括号中的两个求和式

$$\sum_{j,k=1}^{m}\frac{\partial^2 f}{\partial u_j \partial u_k}\mathrm{d}u_j ? \mathrm{d}u_k, \quad \sum_{k,j=1}^{m}\frac{\partial^2 f}{\partial u_k \partial u_j}\mathrm{d}u_k ? \mathrm{d}u_j$$

实际上是相等的,$\mathrm{d}(\mathrm{d}f)=0$,从而对 $\omega\in A^{\circ}(M)$,有 $\mathrm{d}^2\omega=0$,即(5)成立.

由$\mathrm{d}(\omega|_U)|_{U\cap W}=\mathrm{d}(\omega|_{U\cap W})|=\mathrm{d}(\omega|_W)|_{U\cap W}$,知(*)式与(**)式定义的微分算子

$$\mathrm{d}: A(M)\rightarrow A(M)$$

在 $U\cap W$ 上是一致的,亦即,d 在光滑流形 M 上是有定义的、惟一确定的.

对于 $\omega\in A^r(M),\omega=f\cdot \mathrm{d}u_1 ?\cdots ? \mathrm{d}u_r,\ f\in C^{\infty}(M)$

$$\begin{aligned}
&\Rightarrow \mathrm{d}\omega = \mathrm{d}(f\cdot \mathrm{d}u_1 ?\cdots ?\mathrm{d}u_r) = \mathrm{d}f\ ?\mathrm{d}u_1 ?\cdots ?\mathrm{d}u_r \equiv \mathrm{d}f\ ?(\mathrm{d}u_1 ?\cdots ?\mathrm{d}u_r)\\
&\Rightarrow \mathrm{d}(\mathrm{d}\omega) = \mathrm{d}\{\mathrm{d}f\ ?\mathrm{d}(\mathrm{d}u_1 ?\cdots ?\mathrm{d}u_r)\}\\
&\qquad = \mathrm{d}(\mathrm{d}f)?(\mathrm{d}u_1 ?\cdots ?\mathrm{d}u_r) - (-1)^1\mathrm{d}f\ ?\mathrm{d}\{(\mathrm{d}u_1 ?\cdots ?\mathrm{d}u_r)\}\\
&\qquad = 0\ ?(\mathrm{d}u_1 ?\cdots ?\mathrm{d}u_r) - (-1)^1\mathrm{d}f\ ?0 = 0
\end{aligned}$$

最后一步是因为 $\mathrm{d}(\mathrm{d}u_1)=\mathrm{d}(1\cdot \mathrm{d}u_1)=\mathrm{d}(1)?\mathrm{d}u_1=0$. 这就证明了 $\forall\,\omega\in A^r(M)$有 $\mathrm{d}^2\omega=0$. 定理得证.

下面给出例子,取 $M=\mathbb{R}^3$ 的特殊情况来解释一下外微分的意义.

例 6.3.1 $M=\mathbb{R}^3$.

在 $M=\mathbb{R}^3$ 时,取局部坐标$\{u_1,u_2,u_3\}$,$u_1=x,u_2=y,u_3=z$,得到 $m=3$ 维光滑流形.

(1) 点 p 的余切空间 $T_p^*=\mathfrak{F}/\sim=\{[f]^*:[f]\in\mathfrak{F}\}$

自然基为$\{(\mathrm{d}u_k)_p,1\leqslant k\leqslant 3\}=\{(\mathrm{d}u_1)_p,(\mathrm{d}u_2)_p,(\mathrm{d}u_3)_p\}$,亦即

$$\{(\mathrm{d}x)_p,(\mathrm{d}y)_p,(\mathrm{d}z)_p\}\equiv\{\mathrm{d}x,\mathrm{d}y,\mathrm{d}z\},$$

f 在点 p 的微分为 $f\in C_p^{\infty}(\mathbb{R}^3)\Rightarrow(\mathrm{d}f)_p\in T_p^*$,亦即

$$(\mathrm{d}f)_p=\sum_{k=1}^{3}a_k(\mathrm{d}u_k)_p=a_1\mathrm{d}x+a_2\mathrm{d}y+a_3\mathrm{d}z.$$

其中 $a_1=\dfrac{\partial f}{\partial x},\ a_2=\dfrac{\partial f}{\partial y},\ a_3=\dfrac{\partial f}{\partial z}\in C_p^{\infty}(\mathbb{R}^3)$.

(2) 点 p 的切空间 $T_p=(T_p^*)^*$

对偶基为 $\left\{\left.\dfrac{\partial}{\partial u_j}\right|_p,1\leqslant j\leqslant 3\right\}=\left\{\left.\dfrac{\partial}{\partial u_1}\right|_p,\left.\dfrac{\partial}{\partial u_2}\right|_p,\left.\dfrac{\partial}{\partial u_3}\right|_p\right\}$,亦即

$$\left\{\left.\frac{\partial}{\partial x}\right|_p,\left.\frac{\partial}{\partial y}\right|_p,\left.\frac{\partial}{\partial z}\right|_p\right\}\equiv\left\{\frac{\partial}{\partial x},\frac{\partial}{\partial y},\frac{\partial}{\partial z}\right\},$$

点 p 的切矢量 $[\gamma]=X_p\in T_p$,亦即

$$X_p=\sum_{j=1}^{3}\xi_j\left.\frac{\partial}{\partial u_j}\right|_p=\xi_1\frac{\partial}{\partial x}+\xi_2\frac{\partial}{\partial y}+\xi_3\frac{\partial}{\partial z},$$

其中 $\xi_1=\dfrac{\mathrm{d}(x\circ\gamma)}{\mathrm{d}t},\xi_2=\dfrac{\mathrm{d}(y\circ\gamma)}{\mathrm{d}t},\xi_3=\dfrac{\mathrm{d}(z\circ\gamma)}{\mathrm{d}t}$,$\gamma=r(t)$为过点 p 的光滑曲线.

(3) 点 p 的 r 阶协变张量空间 $T^r(T_p^*)\equiv\underbrace{(T_p^*)\otimes\cdots\otimes T_p^*}_{r}$，$0\leqslant r\leqslant 3$.

(4) 点 p 的 r 次外形式空间 $\Lambda^r(T_p^*)\equiv A_r(T^r(T_p^*))$，$0\leqslant r\leqslant 3$

基为$\{(\mathrm{d}u_{k_1})_p ?\cdots?(\mathrm{d}u_{k_r})_p : 1\leqslant k_1,\cdots,k_r\leqslant 3\}$，具体写为

① $\Lambda^0(T_p^*)$的基为$\{1\}$；

② $\Lambda^1(T_p^*)$的基为$\{\mathrm{d}u_1,\mathrm{d}u_2,\mathrm{d}u_3\}$；

③ $\Lambda^2(T_p^*)$的基为$\{\mathrm{d}u_1?\mathrm{d}u_2,\mathrm{d}u_2?\mathrm{d}u_3,\mathrm{d}u_3 ?\mathrm{d}u_1\}$；

④ $\Lambda^3(T_p^*)$的基为$\{\mathrm{d}u_1?\mathrm{d}u_2?\mathrm{d}u_3\}$.

(5) r 次外形式丛 $\Lambda^r((\mathbb{R}^3)^*)=\bigcup\limits_{p\in\mathbb{R}^3}\Lambda^r(T_p^*)$，$0\leqslant r\leqslant 3$.

(6) 外形式丛(r 次外形式丛的分层和)$\Lambda(\mathbb{R}^3)=\sum\limits_{r=0}^{3}\Lambda^r((\mathbb{R}^3)^*)=\sum\limits_{r=0}^{3}\left(\bigcup\limits_{p\in\mathbb{R}^3}\Lambda^r(T_p^*)\right)$.

(7) 外微分式空间(外形式丛的光滑截面空间)$A(\mathbb{R}^3)=\Gamma(\Lambda(\mathbb{R}^3))=\Gamma\left(\sum\limits_{r=0}^{3}\bigcup\limits_{p\in\mathbb{R}^3}\Lambda^r(T_p^*)\right)$

$A(\mathbb{R}^3)$的基$\{(\mathrm{d}u_{k_1})_p ?\cdots?(\mathrm{d}u_{k_r})_p : 1\leqslant k_1,\cdots,k_r\leqslant 3\}$成为

$$\{\{1\},\{\mathrm{d}x,\mathrm{d}y,\mathrm{d}z\},\{\mathrm{d}x ?\mathrm{d}y,\mathrm{d}y?\mathrm{d}z,\mathrm{d}z?\mathrm{d}x\},\{\mathrm{d}x ?\mathrm{d}y?\mathrm{d}z\}\}.$$

于是，$\forall\,\omega\in A(\mathbb{R}^3)$，可以定义一个外微分运算 $\mathrm{d}:A(\mathbb{R}^3)\to A(\mathbb{R}^3)$.

定义一次外微分式为

$$P\,\mathrm{d}x+Q\,\mathrm{d}y+R\,\mathrm{d}z,$$

其中 $P,Q,R\in A^0(\mathbb{R}^3)\equiv C_p^\infty(\mathbb{R}^3)$，则 $\omega\in A^0(\mathbb{R}^3)\Rightarrow \mathrm{d}\omega\in A^1(\mathbb{R}^3)$.

定义二次外微分式为

$$P\,\mathrm{d}x ?\mathrm{d}y+Q\,\mathrm{d}y?\mathrm{d}z+R\,\mathrm{d}z?\mathrm{d}x,$$

其中 $P,Q,R\in A^0(\mathbb{R}^3)$，则 $\omega\in A^1(\mathbb{R}^3)\Rightarrow \mathrm{d}\omega\in A^2(\mathbb{R}^3)$.

定义三次外微分式为

$$P\,\mathrm{d}x ?\mathrm{d}y?\mathrm{d}z,$$

其中 $P\in A^0(\mathbb{R}^3)$，则 $\omega\in A^1(\mathbb{R}^3)\Rightarrow \mathrm{d}\omega\in A^2(\mathbb{R}^3)$；$\omega\in A^2(\mathbb{R}^3)=\mathrm{d}\omega\in A^3(\mathbb{R}^3)$.

不难验证，上面定义的外微分式与外微分满足

① $\mathrm{d}(A^r(M))\subset A^{r+1}(M)$，$r=0,1,2,3$；

② $\omega_1,\omega_2\in A(M)\Rightarrow \mathrm{d}(\omega_1+\omega_2)=\mathrm{d}\omega_1+\mathrm{d}\omega_2$；

③ ω_1 是 r 次外微分式 $\Rightarrow \mathrm{d}(\omega_1?\omega_2)=(\mathrm{d}\omega_1)?\omega_2+(-1)^r\omega_1?\mathrm{d}\omega_2$；

④ f 是M 上的光滑函数，$f\in A^0(M)\Rightarrow \mathrm{d}f$ 恰是 f 在经典意义下的微分；

⑤ $\omega\in A^0(M)\Rightarrow \mathrm{d}(\mathrm{d}\omega)=0$；

⑥ $\mathrm{d}x ?\mathrm{d}x=0$，$\mathrm{d}x ?\mathrm{d}y=-\mathrm{d}y?\mathrm{d}x$，$\mathrm{d}y?\mathrm{d}z=-\mathrm{d}z?\mathrm{d}y$，$\mathrm{d}z?\mathrm{d}x=-\mathrm{d}x ?\mathrm{d}z$.

例 6.3.2　外微分式空间 $A(\mathbb{R}^3)$上的外微分的计算.

$$\forall\,\omega\in A(\mathbb{R}^3)=\Gamma(\Lambda((\mathbb{R}^3)^*))\Rightarrow \omega(p)\in\Lambda((\mathbb{R}^3)^*)=\sum_{r=0}^{3}\Lambda^r((\mathbb{R}^3)^*)$$

$$\Rightarrow \omega = \omega^0 + \omega^1 + \omega^2 + \omega^3,\ \omega^j \in \Lambda^r((\mathbb{R}^3)^*),\quad j = 0,1,2,3.$$

(1) $\omega^0 \in \Lambda^0((\mathbb{R}^3)^*) \Rightarrow \omega = f \in C_p^\infty(\mathbb{R}^3) \Rightarrow$

$$\mathrm{d}\omega^0 = (\mathrm{d}f)_p = P\mathrm{d}x + Q\mathrm{d}y + R\mathrm{d}z = \left(\frac{\partial f}{\partial x}\right)_p \mathrm{d}x + \left(\frac{\partial f}{\partial y}\right)_p \mathrm{d}y + \left(\frac{\partial f}{\partial z}\right)_p \mathrm{d}z,$$

其中$\left(\frac{\partial f}{\partial x}\right)_p, \left(\frac{\partial f}{\partial x}\right)_p, \left(\frac{\partial f}{\partial x}\right)_p \in A^0(\mathbb{R}^3) \equiv C_p^\infty(\mathbb{R}^3)$，则 $\omega^0 \in A^0(\mathbb{R}^3) \Rightarrow \mathrm{d}\omega^0 \in A^1(\mathbb{R}^3) \Rightarrow$

$$\mathrm{d}\omega^0 = (\mathrm{d}f)_p = \left(\frac{\partial f}{\partial x}\right)_p \mathrm{d}x + \left(\frac{\partial f}{\partial y}\right)_p \mathrm{d}y + \left(\frac{\partial f}{\partial z}\right)_p \mathrm{d}z = \mathrm{grad} f \cdot (\mathrm{d}x, \mathrm{d}y, \mathrm{d}z)$$

从而得到光滑函数 $f \in A^0(\mathbb{R}^3)$的外微分 $\mathrm{d}f \in A^1(\mathbb{R}^3)$关于基向量$(\mathrm{d}x,\mathrm{d}y,\mathrm{d}z)$ 的系数是$\mathrm{grad} f$.

(2) $\omega^1 \in \Lambda^1((\mathbb{R}^3)^*) \Rightarrow \omega^1 = a\mathrm{d}x + b\mathrm{d}y + c\mathrm{d}z \Rightarrow$

$$\begin{aligned}\mathrm{d}\omega^1 &= \mathrm{d}(a\mathrm{d}x + b\mathrm{d}y + c\mathrm{d}z) = \mathrm{d}a\ ?\mathrm{d}x + \mathrm{d}b\ ?\mathrm{d}y + \mathrm{d}c\ ?\mathrm{d}z \\ &= \left(\frac{\partial c}{\partial y} - \frac{\partial b}{\partial z}\right)\mathrm{d}y ? \mathrm{d}z + \left(\frac{\partial a}{\partial z} - \frac{\partial c}{\partial x}\right)\mathrm{d}z ? \mathrm{d}x + \left(\frac{\partial b}{\partial x} - \frac{\partial a}{\partial y}\right)\mathrm{d}x\ ?\mathrm{d}y,\end{aligned}$$

其中 $a,b,c \in A^0(\mathbb{R}^3)$，则 $\omega^1 \in A^1(\mathbb{R}^3) \Rightarrow \mathrm{d}\omega^1 \in A^2(\mathbb{R}^3) \Rightarrow$

$$\mathrm{d}\omega^1 = \mathrm{rot}(a,b,c) \cdot (\mathrm{d}y ? \mathrm{d}z, \mathrm{d}z ? \mathrm{d}x, \mathrm{d}x\ ?\mathrm{d}y)$$

从而得到一次外微分式 $\omega^1 = (a,b,c) \in A^1(\mathbb{R}^3)$的外微分 $\mathrm{d}\omega^1 \in A^2(\mathbb{R}^3)$关于基向量$(\mathrm{d}y ? \mathrm{d}z, \mathrm{d}z ? \mathrm{d}x, \mathrm{d}x\ ?\mathrm{d}y)$的系数是 $\mathrm{rot}(a,b,c)$.

(3) $\omega^2 \in \Lambda^2((\mathbb{R}^3)^*) \Rightarrow \omega^2 = a\mathrm{d}y ? \mathrm{d}z + b\mathrm{d}z ? \mathrm{d}x + c\mathrm{d}x\ ?\mathrm{d}y \Rightarrow$

$$\begin{aligned}\mathrm{d}\omega^2 &= \mathrm{d}(a\mathrm{d}y ? \mathrm{d}z + b\mathrm{d}z ? \mathrm{d}x + c\mathrm{d}x\ ?\mathrm{d}y) \\ &= \mathrm{d}a\ ?\mathrm{d}y ? \mathrm{d}z + \mathrm{d}b\ ?\mathrm{d}z ? \mathrm{d}x + \mathrm{d}c\ ?\mathrm{d}x\ ?\mathrm{d}y \\ &= \left(\frac{\partial a}{\partial x}\mathrm{d}x + \frac{\partial a}{\partial y}\mathrm{d}y + \frac{\partial a}{\partial z}\mathrm{d}z\right) ? \mathrm{d}y ? \mathrm{d}z + \left(\frac{\partial b}{\partial x}\mathrm{d}x + \frac{\partial b}{\partial y}\mathrm{d}y + \frac{\partial b}{\partial z}\mathrm{d}z\right) ? \mathrm{d}z ? \mathrm{d}x \\ &\quad + \left(\frac{\partial c}{\partial x}\mathrm{d}x + \frac{\partial c}{\partial y}\mathrm{d}y + \frac{\partial c}{\partial z}\mathrm{d}z\right) ? \mathrm{d}x\ ?\mathrm{d}y \\ &= \frac{\partial a}{\partial x}\mathrm{d}x\ ?\mathrm{d}y ? \mathrm{d}z + \frac{\partial b}{\partial y}\mathrm{d}y ? \mathrm{d}z ? \mathrm{d}x + \frac{\partial c}{\partial z}\mathrm{d}z ? \mathrm{d}x\ ?\mathrm{d}y \\ &= \frac{\partial a}{\partial x}\mathrm{d}x\ ?\mathrm{d}y ? \mathrm{d}z - \frac{\partial b}{\partial y}\mathrm{d}y ? \mathrm{d}x\ ?\mathrm{d}z - \frac{\partial c}{\partial z}\mathrm{d}x\ ?\mathrm{d}z ? \mathrm{d}y \\ &= \frac{\partial a}{\partial x}\mathrm{d}x\ ?\mathrm{d}y ? \mathrm{d}z - (-1)\frac{\partial b}{\partial y}\mathrm{d}x\ ?\mathrm{d}y ? \mathrm{d}z - (-1)\frac{\partial c}{\partial z}\mathrm{d}x\ ?\mathrm{d}y ? \mathrm{d}z \\ &= \left(\frac{\partial a}{\partial x} + \frac{\partial b}{\partial y} + \frac{\partial c}{\partial z}\right)\mathrm{d}x\ ?\mathrm{d}y ? \mathrm{d}z,\end{aligned}$$

其中 $a,b,c \in A^0(\mathbb{R}^3)$，则 $\omega^2 \in A^2(\mathbb{R}^3) \Rightarrow \mathrm{d}\omega^2 \in A^3(\mathbb{R}^3) \Rightarrow$

$$\mathrm{d}\omega^2 = \left(\frac{\partial a}{\partial x} + \frac{\partial b}{\partial y} + \frac{\partial c}{\partial z}\right)\mathrm{d}x\ ?\mathrm{d}y ? \mathrm{d}z = \mathrm{div}(a,b,c)(\mathrm{d}x\ ?\mathrm{d}y ? \mathrm{d}z)$$

从而得到二次外微分式 $\omega^2 = (a,b,c) \in A^2(\mathbb{R}^3)$的外微分为 $\mathrm{d}\omega^2 \in A^3(\mathbb{R}^3)$关于基向量$(\mathrm{d}x\ ?\mathrm{d}y ? \mathrm{d}z)$的系数是 $\mathrm{div}(a,b,c)$.

(4) $\omega^3 \in \Lambda^3((\mathbb{R}^3)^*) \Rightarrow \omega^3 = a\mathrm{d}x ? \mathrm{d}y ? \mathrm{d}z \Rightarrow$ 其中 $a \in A^0(\mathbb{R}^3) \Rightarrow \omega^3 \in A^3(\mathbb{R}^3) \Rightarrow \mathrm{d}\omega^3 = 0$.

3. 外微分式空间 $A(M)$ 上的外微分的性质

定理 6.3.2(Poincare 引理) 设 M 是 m 维光滑流形,则对任意外微分式 $\omega \in A(M)$,其二次外微分 $\mathrm{d}^2\omega = \mathrm{d}(\mathrm{d}\omega)$ 都有 $\mathrm{d}^2\omega = 0$.

证 因为外微分算子 d 是线性的,故只要对 $\omega \in A(M)$ 是单项式的情形加以证明. 由于外微分算子的局部性,只需假设 $\omega = a(\mathrm{d}u_1 ? \mathrm{d}u_2 ? \cdots ? \mathrm{d}u_r)$. 于是,$\mathrm{d}\omega = \mathrm{d}a ? \mathrm{d}u_1 ? \mathrm{d}u_2 ? \cdots ? \mathrm{d}u_r$,故得到 $\mathrm{d}^2\omega = \mathrm{d}(\mathrm{d}a ? \mathrm{d}u_1 ? \mathrm{d}u_2 ? \cdots ? \mathrm{d}u_r)$,利用外微分的定义中的(3),$\omega_1$ 是 r 次外微分式蕴含 $\mathrm{d}(\omega_1 ? \omega_2) = (\mathrm{d}\omega_1) ? \omega_2 + (-1)^r \omega_1 ? \mathrm{d}\omega_2$,与(5)得到 $\omega \in A^0(M) \Rightarrow \mathrm{d}(\mathrm{d}\omega) = 0$,故

$$\mathrm{d}^2\omega = \mathrm{d}(\mathrm{d}a) ? (\mathrm{d}u_1 ? \mathrm{d}u_2 ? \cdots ? \mathrm{d}u_r) - \mathrm{d}a ? \mathrm{d}(\mathrm{d}u_1) ? \cdots ? \mathrm{d}u_r + \cdots = 0,$$

定理得证.

由定理 6.3.2,立即可以得到经典微积分中的关于场论的基本公式:

$$\mathrm{rot}(\mathrm{grad} f) = 0, \quad \mathrm{div}(\mathrm{rot}(a, b, c)) = 0$$

定理 6.3.3 设 ω 是光滑流形 M 上的一次外微分式,X, Y 是 M 上的光滑切矢量场,则

$$\langle X ? Y, \mathrm{d}\omega \rangle = X\langle Y, \omega \rangle - Y\langle X, \omega \rangle - \langle [X, Y], \omega \rangle,$$

其中$[X,Y] = XY - YX : C^\infty(M) \to C^\infty(M)$为 Poisson 括号

$$[X, Y](f) = X(Yf) - Y(Xf), \quad f \in C^\infty(M).$$

定理 6.3.4 设 $F: M \to N$ 是光滑流形 M 到 N 的光滑映射,则由 F 诱导的线性映射

$$F^* : A(N) \to A(M)$$

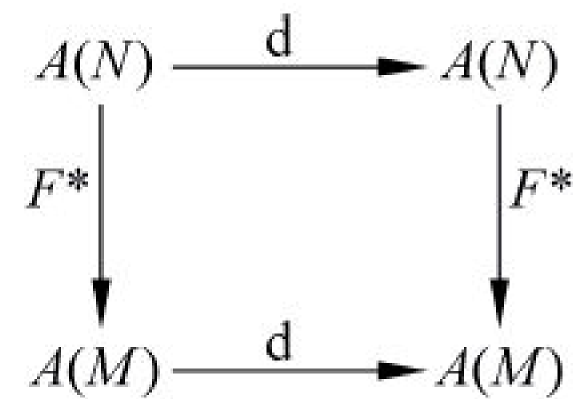


图 6.3.5 交换图

是外微分式空间 $A(N)$ 到外微分式空间 $A(M)$ 的光滑映射,且 F^* 与外微分算子 d 是可交换的,$F^* \circ \mathrm{d} = \mathrm{d} \circ F^* : A(N) \to A(M)$,亦即交换图 6.3.5 成立,这里映射 F^* 的定义见(6.2.27)式.

例 6.3.3 设 $\omega = xy\mathrm{d}x + z\mathrm{d}y - yz\mathrm{d}z$,光滑映射 $F: \mathbb{R}^2 \to \mathbb{R}^3$,由 $F(u,v) = (uv, u^2, 3u+v)$ 给出,求 $F^*\omega$ 与 $F^*\mathrm{d}\omega$.

解 先求 $\mathrm{d}\omega$,有

$$\begin{aligned}\mathrm{d}\omega &= (x\mathrm{d}y + y\mathrm{d}x) ? \mathrm{d}x + \mathrm{d}z ? \mathrm{d}z - (y\mathrm{d}x + z\mathrm{d}y) ? \mathrm{d}z \\ &= -x\mathrm{d}x ? \mathrm{d}y - (1+z)\mathrm{d}y ? \mathrm{d}z.\end{aligned}$$

为求 $F^*\mathrm{d}x$,只需由 $F(u,v) = (uv, u^2, 3u+v)$确定的 $x = uv, y = u^2, z = 3u+v$,将 $F^*\mathrm{d}x$ 中出现的 x, y, z 及所有 $\mathrm{d}x, \mathrm{d}y, \mathrm{d}z$ 通过映射 $F: \mathbb{R}^2 \to \mathbb{R}^3$ 换为 u, v, w 与 $\mathrm{d}u, \mathrm{d}v, \mathrm{d}w$,于是,$F^*\mathrm{d}x = v\mathrm{d}u + u\mathrm{d}v$; 其余 $F^*\mathrm{d}y, F^*\mathrm{d}z$ 类推. 故有

$$F^*\mathrm{d}x = v\mathrm{d}u + u\mathrm{d}v, \quad F^*\mathrm{d}y = 2u\mathrm{d}u, \quad F^*\mathrm{d}z = 3\mathrm{d}u + \mathrm{d}v,$$

$$F^*(\mathrm{d}x ? \mathrm{d}y) = F^*\mathrm{d}x ? F^*\mathrm{d}y = (v\mathrm{d}u + u\mathrm{d}v) ? (2u\mathrm{d}u) = -2u^2\mathrm{d}u ? \mathrm{d}v,$$

$$F^*(\mathrm{d}y ? \mathrm{d}z) = F^*\mathrm{d}y ? F^*\mathrm{d}z = (2u\mathrm{d}u) ? (3\mathrm{d}u + \mathrm{d}v) = 2u\mathrm{d}u ? \mathrm{d}v,$$

$$F^*(\mathrm{d}z ? \mathrm{d}x) = F^*\mathrm{d}z ? F^*\mathrm{d}x = (3\mathrm{d}u + \mathrm{d}v) ? (v\mathrm{d}u + u\mathrm{d}v) = (3u - v)\mathrm{d}u ? \mathrm{d}v.$$

进而，由 $\omega = xy\mathrm{d}x + z\mathrm{d}y - yz\mathrm{d}z$，故

$$\begin{aligned} F^*\omega &= (xy\mathrm{d}x + z\mathrm{d}y - yz\mathrm{d}z)\Big|_{x=uv,y=u^2,z=3u+v} \\ &= u^3v(v\mathrm{d}u + u\mathrm{d}v) + 2u(3u+v)\mathrm{d}u - u^2(3u+v)(3\mathrm{d}u + \mathrm{d}v), \end{aligned}$$

化简后为

$$F^*\omega = (u^3v^2 + 6u^2 + 2uv - 9u^3 - 3u^2v)\mathrm{d}u + (u^4v - 3u^3 - u^2v)\mathrm{d}v,$$

类似地，有

$$\begin{aligned} F^*\mathrm{d}\omega &= -uv(-2u^2\mathrm{d}u ? \mathrm{d}v) - 2(1 + 3u + v)u\mathrm{d}u ? \mathrm{d}v \\ &= (2u^3v - 6u^2 - 2uv - 2u)\mathrm{d}u ? \mathrm{d}v. \end{aligned}$$

6.4 外微分式的积分

在给出流形上的外微分式及其外微分(局部性质)之后，我们还要研究外微分式的积分(整体性质). 为将流形的局部性质用到整体性质的研究中去，最重要的方式就是考虑外微分式在流形上的积分.

6.4.1 光滑流形的定向

1. 光滑流形 M 的定向

定义 6.4.1(定向流形) m 维光滑流形 $M \equiv (M, \mathbb{A})$ 称为**可定向的**，若存在 M 上的一个连续的、处处不为零的 m 次外微分式 $\omega \in A^m(M)$.

若在 M 上给定了这样的非零 m 次外微分式 $\omega \in A^m(M)$，$\omega \neq 0$，则称 M 是**定向流形**(**directed manifold**).

注 关于流形的定向有以下注记.

(1) 用局部坐标系来刻画定向——对于光滑流形 M，设 $\Sigma_0 = \{(U_\alpha, x_\alpha^j) : \alpha \in I, 1 \leqslant j \leqslant m\}$

是一个坐标开覆盖，使得 $\forall\alpha,\beta\in I$，且 $U_\alpha\cap U_\beta\neq\varnothing$ 时，坐标变换的 Jacobi 行列式在 $U_\alpha\cap U_\beta$ 上处处保持定号，例如 $\det\left[\dfrac{\partial x_\alpha^j}{\partial x_\beta^k}\right]\Bigg|_{U_\alpha\cap U_\beta}$ 恒为正，或恒为负，则 M 是一个定向的光滑流形，称 Σ_0 为光滑流形 M 的定向光滑坐标覆盖系；

(2) 若 M 是满足第二可数公理的连通光滑流形，则 M 是可定向流形，当且仅当 M 上存在一个处处不为零的 m 次外微分式 $\omega\in A^m(M)$. 显然，可以确定 M 定向的外微分式 $\omega\in A^m(M)$，$\omega\neq 0$ 不是惟一的；

(3) 若 M 是定向连通光滑流形，当确定 M 定向的两个外微分式彼此间相差一个处处为正的函数因子，则称这两个外微分式给定了流形 M 的同一个定向. 这是因为，设 $\omega,\omega'\in A^m(M)$是给出流形定向的两个 m 次外微分式，则存在 M 上的处处不为零的连续函数 f，使得 $\omega'=f\omega$. 由于 M 是连通的，这个 f 必须在 M 上保持同一个符号，故 $\omega'\in A^m(M)$给出的定向，或者与 $\omega\in A^m(M)$给出的定向一致，或者与$(-\omega)\in A^m(M)$给出的定向一致. 于是，$f>0$ 时，ω 与 ω' 确定 M 的同一个定向. M 恰好可以有两个不同的定向.

与$\mathbb{R}^3$ 情形一致，空间的双侧曲面是定向曲面，恰好有两个不同的定向.

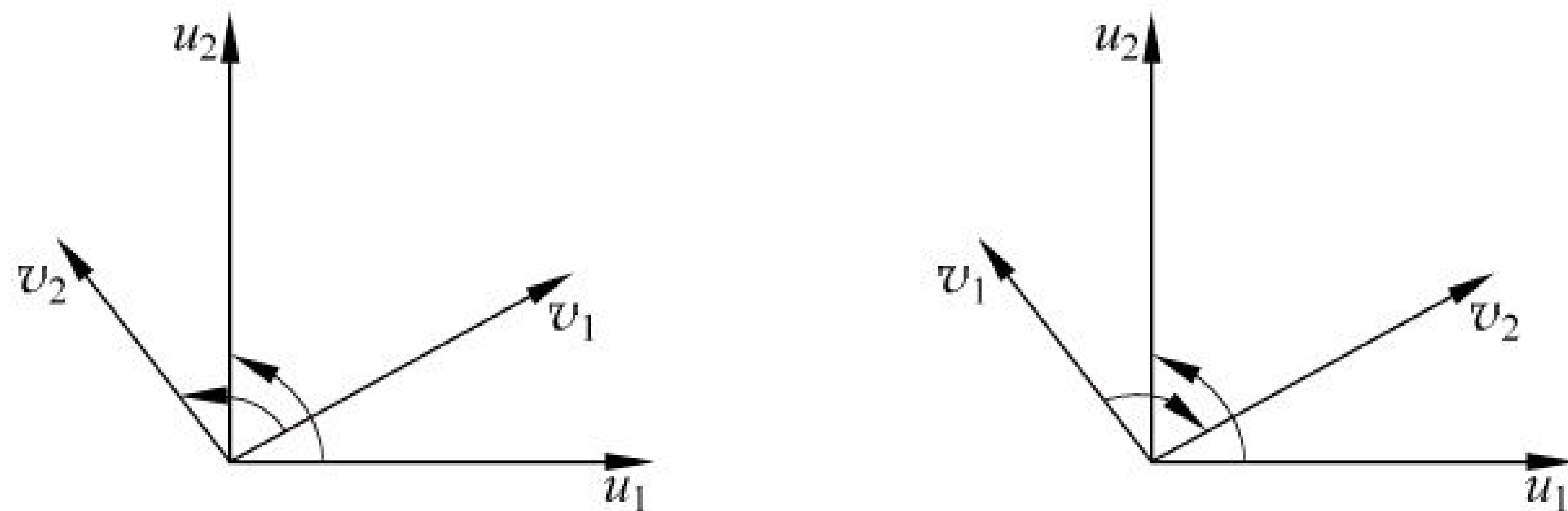
例 6.4.1 光滑流形 $M=\mathbb{R}^1,\mathbb{R}^2,\mathbb{R}^3$.

对于 $M=\mathbb{R}^1,\omega\in A^1(\mathbb{R}^1)\Rightarrow\omega=a\mathrm{d}x,a\in C^\infty(\mathbb{R}^3)\Rightarrow$ 取两个坐标系$\{u\}$、$\{v\}$，则有下面两个方向：

du ⟶ dv　　　　dv ⟵⟶ du

$\mathbb{R}^1$ 的定向：取 $A(\mathbb{R}^1)$的基$\{\{1\};\{\mathrm{d}x\}\}$，则$\{\mathrm{d}x\}$是 $A^1(\mathbb{R}^1)$的基，能表示$\mathbb{R}^1$ 的方向. 显然，$\exists\omega=a\mathrm{d}x\in A^1(\mathbb{R}^1)$，$a\in C^\infty(\mathbb{R}^1)$，确定 $\mathbb{R}^1$ 的方向. 例如取 $a=1$，则 $\omega=a\mathrm{d}x=\mathrm{d}x$，$A^1(\mathbb{R}^1)$的两个基之间有一个变换方阵，变换方阵的行列式为正，则两个基之间确定同一定向；若变换方阵的行列式为负，则两个基确定不同的方向.

对于 $M=\mathbb{R}^2,\omega\in A^2(\mathbb{R}^2)\Rightarrow\omega=a\mathrm{d}x$?$\mathrm{d}y,a\in C^\infty(\mathbb{R}^3)\Rightarrow$ 取两个坐标系$\{u_1,u_2\}$，$\{v_1,v_2\}$，则有下面两个方向：



$\mathbb{R}^2$ 的定向：取 $A(\mathbb{R}^2)$的基$\{1;\{\mathrm{d}x,\mathrm{d}y\};\{\mathrm{d}x$?$\mathrm{d}y\}\}$，则由 $A^2(\mathbb{R}^2)$的基$\{\mathrm{d}x$?$\mathrm{d}y\}$(一定要取 $A^2(\mathbb{R}^2)$的基，才能表示 $\mathbb{R}^2$ 的方向)，$\exists\omega=a\mathrm{d}x$?$\mathrm{d}y,\omega\neq 0,a\in C^\infty(\mathbb{R}^2)$，确定$\mathbb{R}^2$ 的方向. 例如，取 $a=1$，由于 $\mathrm{d}x$?$\mathrm{d}y=-\mathrm{d}y$?$\mathrm{d}x$，因此取 $\mathrm{d}x$?$\mathrm{d}y$ 与 $\mathrm{d}y$?$\mathrm{d}x$ 正好确定两个不同的方向.

对于 $M=\mathbb{R}^3,\omega\in A^3(\mathbb{R}^3)\Rightarrow\omega=a\mathrm{d}x$?$\mathrm{d}y$?$\mathrm{d}z,a\in C^\infty(\mathbb{R}^3)$.

$\mathbb{R}^3$ 的定向：取 $A(\mathbb{R}^3)$的基$\{1;\{\mathrm{d}x,\mathrm{d}y,\mathrm{d}z\};\{\mathrm{d}y?\mathrm{d}z,\mathrm{d}z?\mathrm{d}x,\mathrm{d}z?\mathrm{d}y\};\{\mathrm{d}x?\mathrm{d}y?\mathrm{d}z\}\}$，则由 $A^3(\mathbb{R}^3)$的基$\{\mathrm{d}x?\mathrm{d}y?\mathrm{d}z\}$（一定要取 $A^3(\mathbb{R}^3)$的基，才能表示$\mathbb{R}^3$ 的方向），$\exists\,\omega=a\mathrm{d}x?\mathrm{d}y?\mathrm{d}z,\omega\neq 0,a\in C^\infty(\mathbb{R}^3)$，确定$\mathbb{R}^3$ 的方向. 若取两个坐标系$\{u_1,u_2,u_3\},\{v_1,v_2,v_3\}$，则 $\mathrm{d}u_1?\mathrm{d}u_2?\mathrm{d}u_3$ 与 $\mathrm{d}v_1?\mathrm{d}v_2?\mathrm{d}v_3$ 只可能有两个方向：

2. 光滑流形 M 上的单位分解定理

为给出定向 $\omega\in A^m(M)$的坐标表示，我们首先给出以下定义.

定义 6.4.2（函数、外微分式的支集） 设 M 为 m 维光滑流形，$f:M\to\mathbb{R}$ 是 M 上的实值函数，函数 f 的支集是指：使得 f 取非零值的点集的闭包，记为

$$\mathrm{supp}f=\overline{\{p\in M:\ f(p)\neq 0\}};$$

函数 f 的支集的补集，恰是流形 M 中使得 $f=0$ 的最大开子集.

若 $\omega\in A(M)$是一个外微分式，则外微分式 ω 的支集定义为

$$\mathrm{supp}\,\omega=\overline{\{p\in M:\ \omega(p)\neq 0\}}.$$

外微分式 ω 的支集的补集，恰是流形 M 中使得 $\omega=0$ 的最大开子集.

定义 6.4.3（流形 M 的局部有限开覆盖） 设 Σ_0 是 M 的一个开覆盖，如果 M 的任意紧致子集 $N\subset M$ 仅与 Σ_0 中有限多个开集相交，则称 Σ_0 是 M 的**局部有限开覆盖**.

定理 6.4.1（局部有限开覆盖的存在性定理） 设 Σ 是 m 维光滑流形 M 的一个拓扑基，则存在 Σ 的一个子集 Σ_0，使得 Σ_0 是 M 的局部有限开覆盖.

证 用归纳法证明，参看[3].

定理 6.4.2（单位分解定理） 设 $\Sigma=\{W_j\}$是 m 维光滑流形 M 的一个开覆盖，则在 M 上存在一组光滑函数$\{g_\alpha\}$，满足下列条件：

（1）对于每个指标 α，都有 ① $0\leqslant g_\alpha\leqslant 1$；② $\mathrm{supp}g_\alpha$ 是 M 中的紧致集；③ 存在开集 $W_j\in\Sigma$，使$\mathrm{supp}g_\alpha\subset W_j$；④ $\sum\limits_\alpha g_\alpha=1$.

（2）对任意 $p\in M$，存在开邻域 U，使得 U 只与有限多个支集$\mathrm{supp}g_\alpha$ 相交.

证 请读者参看[3].

注 由条件(2)知,④ 中的和式仅仅是有限和. 我们称$\{g_\alpha\}$为**从属于开覆盖 Σ 的单位分解**,记为

$$(\Sigma=\{W_j\},g=\{g_\alpha\}).$$

单位分解实际上是指"1 的分解",因为 $1=\sum\limits_\alpha g_\alpha$ 可理解为 1 被"分解"了.

6.4.2 外微分式在定向光滑流形上的积分

1. m 次外微分式的积分的定义

设 M 为 m 维定向光滑流形,$\omega\in A^m(M)$是 M 上的 m 次外微分式,且 ω 具有紧致支集,即$\mathrm{supp}\omega$ 是紧致集. 又设 $\Sigma=\{W_j\}$是与 M 的一个定向相符合的坐标覆盖,$\{g_\alpha\}$是从属于 Σ 的单位分解,则

$$\omega=\left(\sum_\alpha g_\alpha\right)\cdot\omega=\sum_\alpha g_\alpha\cdot\omega,$$

上式求和是有限和,对于和式的每一项 $g_\alpha\cdot\omega$,由于$\mathrm{supp}\omega$ 是紧致集,故对$\mathrm{supp}g_\alpha\cdot\omega$,存在 $W_j\in\Sigma$,使 $\mathrm{supp}(g_\alpha\cdot\omega)\subset W_j$. 于是,我们可定义 $g_\alpha\cdot\omega$ 在 M 上的积分为

$$\int_{W_j}g_\alpha\cdot\omega,\tag{6.4.1}$$

上式是通常意义下的 Riemann 积分,亦即,将 $g_\alpha\cdot\omega$ 表示为关于 W_j 中的局部坐标 $u_1,\cdots,u_m$ 的表示 $g_\alpha\cdot\omega\,\mathrm{d}u_1$?…?$\mathrm{d}u_m=f(u_1,\cdots,u_m)\mathrm{d}u_1\cdots\mathrm{d}u_m$,则右边积分是在 $\mathbb{R}^m$ 中的开集 W_j 上的 m 重 Riemann 积分

$$\int_{W_j}g_\alpha\cdot\omega\equiv\int_{W_j}f(u_1,\cdots,u_m)\mathrm{d}u_1\cdots\mathrm{d}u_m.\tag{6.4.2}$$

注 在上述积分定义中,必须说明定义有意义. 为此,需说明两件事:

① 在给定 $(\Sigma=\{W_j\},g=\{g_\alpha\})$ 后,(6.4.2) 式中的$\int_{W_j}g_\alpha\cdot\omega$ 与$\int_{W_k}g_\alpha\cdot\omega$ 在 $W_j\cap W_k$,$j\neq k$ 时,其值相等,即$\int_{W_j}g_\alpha\cdot\omega=\int_{W_k}g_\alpha\cdot\omega$,亦即,其值与局部坐标系的选取无关.

事实上,设 $\mathrm{supp}(g_\alpha\cdot\omega)\subset W_j$ 与 $\mathrm{supp}(g_\alpha\cdot\omega)\subset W_k$ 同时成立,且

$$W_j\leftrightarrow u_1,\cdots,u_m,\quad W_k\leftrightarrow v_1,\cdots,v_m,$$

则坐标变换的 Jacobi 行列式为 $J=\dfrac{\partial(v_1,\cdots,v_m)}{\partial(u_1,\cdots,u_m)}$.

设 $J=\dfrac{\partial(v_1,\cdots,v_m)}{\partial(u_1,\cdots,u_m)}>0$($J<0$ 类似讨论),且 $g_\alpha\cdot\omega$ 表示为关于 W_j,W_k 中局部坐标表示式分别为

$$g_\alpha\cdot\omega\,\mathrm{d}u_1?\cdots?\mathrm{d}u_m\equiv f(u_1,\cdots,u_m)\mathrm{d}u_1\cdots\mathrm{d}u_m$$

与

$$g_\alpha \cdot \omega \, \mathrm{d}v_1 ?\cdots ? \mathrm{d}v_m \equiv h(v_1, \cdots, v_m) \mathrm{d}v_1 \cdots \mathrm{d}v_m,$$

则 $f(u_1,\cdots,u_m)=h(v_1,\cdots,v_m)\cdot|J|$. 根据 Riemann 积分的积分变量代换公式,有

$$\int_{W_j \cap W_k} g_\alpha \cdot \omega = \int_{W_j \cap W_k} h(v_1, \cdots, v_m) \mathrm{d}v_1 \cdots \mathrm{d}v_m$$

$$= \int_{W_j \cap W_k} h(v_1, \cdots, v_m) |J| \mathrm{d}u_1 \cdots \mathrm{d}u_m = \int_{W_j \cap W_k} f(u_1, \cdots, u_m) \mathrm{d}u_1 \cdots \mathrm{d}u_m,$$

因此,(6.4.2)式与坐标系的选取无关.

进而,由于suppω 是紧致集,由从属于开覆盖 $\Sigma=\{W_j\}$的单位分解$\{g_\alpha\}$的条件②,知与 suppω 相交的 W_j 只有有限多个,因此,可定义

$$\int_M \omega = \int_M \sum_\alpha g_\alpha \cdot \omega = \sum_\alpha \int_M g_\alpha \cdot \omega, \tag{6.4.3}$$

并且以上推理说明:对于每一对$(\Sigma=\{W_j\}, g=\{g_\alpha\})$,(6.4.3)式是惟一确定的.

② 对于 $\Sigma=\{W_j\}$的不同的单位分解$(\Sigma=\{W_j\}, g=\{g_\alpha\})$,$(\Sigma=\{W_j\}, g'=\{g'_\beta\})$,(6.4.3)式右边的值相等.

事实上,将 g'_β 在 $g=\{g_\alpha\}$之下进行分解,$g'_\beta = \sum\limits_\alpha g_\alpha \cdot g'_\beta$,于是

$$\sum_\beta \int_M g'_\beta \cdot \omega = \sum_\beta \int_M g'_\beta \left(\sum_\alpha g_\alpha \cdot \omega \right) = \sum_\beta \sum_\alpha \int_M g'_\beta \cdot g_\alpha \cdot \omega$$

$$= \sum_\alpha \sum_\beta \int_M g_\alpha \cdot g'_\beta \cdot \omega = \sum_\alpha \int_M g_\alpha \left(\sum_\beta g'_\beta \cdot \omega \right) = \sum_\alpha \int_M g_\alpha \cdot \omega.$$

现在可以给出 m 维光滑流形 M 上的 m 次外微分式的积分的定义了.

定义 6.4.4(m 次外微分式的积分) 光滑流形 M 上的、具有紧致支集的 m 次外微分式 $\omega \in A^m(M)$在 M 上的积分定义为

$$\int_M \omega = \sum_\alpha \int_M g_\alpha \cdot \omega, \tag{6.4.4}$$

(6.4.4)式的值与$(\Sigma=\{W_j\}, g=\{g_\alpha\})$的选取无关.

定理 6.4.3(m 次外微分式积分的性质) 对于光滑流形 M 上的、具有紧致支集的 m 次外微分式 $\omega_1, \omega_2 \in A^m(M)$与任意实数 $c_1, c_2 \in \mathbb{R}$,在 M 上的积分满足**线性性质**

$$\int_M c_1\omega_1 + c_2\omega_2 = c_1 \int_M \omega_1 + c_2 \int_M \omega_2.$$

2. r 次($r<m$)外微分式的积分的定义

若 $\omega \in A^r(M)$,$r<m$,具有紧致支集 supp $\omega \subset N$,这里 N 是 M 的 r 维子流形. 假设映射 $h: N \to M$ 是 N 到 M 的一个**嵌入映射**,使得 N 是 M 的 r 维嵌入子流形,则 $h: N \to M$ 的诱导出的映射 $h^*: A(M) \to A(N)$使得:$\omega \in A^r(M)(r<m)$蕴含 $h^*\omega \in A^r(N)$,且 $h^*\omega$ 是 r 维光滑流形 N 上的 r 次外微分式,具有紧致支集 supp$h^*\omega \subset N$. 故积分

$$\int_M \omega = \int_{h(N)} \omega = \int_N h^* \omega$$

有意义，从而低维外微分式 $\omega \in A^r(M)(r < m)$ 的积分可以定义如下.

定义 6.4.5（r **次外微分式**（$r<m$）**的积分**） 光滑流形 M 上的具有紧致支集的 r 次外微分式 $\omega \in A^r(M)$，$r<m$，在 M 上的积分定义为

$$\int_M \omega = \int_{h(N)} \omega = \int_N h^* \omega. \tag{6.4.5}$$

(6.4.5)式用低维外微分式 $\omega \in A^r(M)$ 在同维空间 N，$\dim N = r < m$ 上的积分（按定义 6.4.4）来定义 $\omega \in A^r(M)$ 在空间 M 上的积分；其中由嵌入映射 $h: N \to M$ 诱导出的映射 $h^*: A(M) \to A(N)$ 由(6.2.27)式给出.

例 6.4.2 外微分式 $\omega \in A^r(M)$ 在光滑流形上的积分在 $M = \mathbb{R}^3$ 时的表示.

① $M = \mathbb{R}^3$ 上的 3 次外微分式 $\omega \in A^3(\mathbb{R}^3)$ 的积分. 取局部坐标 $\{u_1, u_2, u_3\}$ 为 $u_1 = x, u_2 = y, u_3 = z$，则有 $A^3(\mathbb{R}^3)$ 的基

$$\{\mathrm{d}u_1 ? \mathrm{d}u_2 ? \mathrm{d}u_3\} = \{\mathrm{d}x, \mathrm{d}y, \mathrm{d}z\},$$

从而 $\omega \in A^3(\mathbb{R}^3) \Rightarrow \omega = H\mathrm{d}x ? \mathrm{d}y ? \mathrm{d}z$，其中 $H \in A^0(\mathbb{R}^3)$，于是

$$\int_{\mathbb{R}^3} \omega = \int_{\mathbb{R}^3} H(u_1, u_2, u_3)\mathrm{d}u_1 \mathrm{d}u_2 \mathrm{d}u_3 = \int_{\mathbb{R}^3} H(x, y, z)\mathrm{d}x\mathrm{d}y\mathrm{d}z$$

是 $\mathbb{R}^3$ 上的三重积分，为与 $\mathbb{R}^3$ 上的通常记号一致，记为

$$\int_{\mathbb{R}^3} \omega \,\mathrm{d}x ? \mathrm{d}y ? \mathrm{d}z = \int_{\mathbb{R}^3} H(x, y, z)\mathrm{d}x\mathrm{d}y\mathrm{d}z.$$

② $M = \mathbb{R}^3$ 上的 2 次外微分式 $\omega \in A^2(\mathbb{R}^3)$ 的积分. 取局部坐标 $\{u_1, u_2, u_3\}$，为 $u_1 = x$，$u_2 = y, u_3 = z$，则有 $N = \mathbb{R}^2$，$A^2(\mathbb{R}^3)$ 的基

$$\{\mathrm{d}y ? \mathrm{d}z, \mathrm{d}z ? \mathrm{d}x, \mathrm{d}x ? \mathrm{d}y\},$$

于是，$\omega \in A^2(\mathbb{R}^3) \Rightarrow \omega = P\mathrm{d}y ? \mathrm{d}z + Q\mathrm{d}z ? \mathrm{d}x + R\mathrm{d}x ? \mathrm{d}y$，其中 $P, Q, R \in A^0(\mathbb{R}^3)$.

取嵌入映射 $h: N \to M$ 为 $I: \mathbb{R}^2 \to \mathbb{R}^3$ 恒同映射，于是

$$\int_{\mathbb{R}^3} \omega = \int_{\mathbb{R}^2} h^* \circ \omega,$$

其中 $h^* \circ \omega \in A^2(\mathbb{R}^2)$ 在所取的局部坐标系 $u_1 = x, u_2 = y, u_3 = z$ 之下，$h^* \circ \omega = h^*(\omega) = I^*(\omega)$，就是 $\omega = P\mathrm{d}y ? \mathrm{d}z + Q\mathrm{d}z ? \mathrm{d}x + R\mathrm{d}x ? \mathrm{d}y$ 在局部坐标系上的限制，$I^*(\omega) = \omega|_U$，但由局部坐标系的取法，知

$$\begin{aligned} I^*(\omega) &= I^*(P\mathrm{d}y ? \mathrm{d}z + Q\mathrm{d}z ? \mathrm{d}x + R\mathrm{d}x ? \mathrm{d}y) \\ &= P\mathrm{d}y ? \mathrm{d}z + Q\mathrm{d}z ? \mathrm{d}x + R\mathrm{d}x ? \mathrm{d}y, \end{aligned}$$

故

$$h^* \circ \omega|_U = P\mathrm{d}y ? \mathrm{d}z + Q\mathrm{d}z ? \mathrm{d}x + R\mathrm{d}x ? \mathrm{d}y,$$

从而

$$\int_{\mathbb{R}^2} h^* \circ \omega = \int_{\mathbb{R}^2} P\mathrm{d}u_2 ? \mathrm{d}u_3 + Q\mathrm{d}u_3 ? \mathrm{d}u_1 + R\mathrm{d}u_1 ? \mathrm{d}u_2$$

$$= \int_{\mathbb{R}^2} P\mathrm{d}y ? \mathrm{d}z + Q\mathrm{d}z ? \mathrm{d}x + R\mathrm{d}x ? \mathrm{d}y$$

$$= \int_{\mathbb{R}^2} P\mathrm{d}y\mathrm{d}z + Q\mathrm{d}z\mathrm{d}x + R\mathrm{d}x\mathrm{d}y.$$

③ $M=\mathbb{R}^3$ 上的 1 次外微分式 $\omega\in A^1(\mathbb{R}^3)$的积分. 取局部坐标$\{u_1,u_2,u_3\}$为 $u_1=x,u_2=y,u_3=z$,则有 $N=\mathbb{R}^1$,$A^1(\mathbb{R}^3)$的基底

$$\{\mathrm{d}x,\mathrm{d}y,\mathrm{d}z\},$$

于是,$\omega\in A^2(\mathbb{R}^3) \Rightarrow \omega=P\mathrm{d}x+Q\mathrm{d}y+R\mathrm{d}z$,其中 $P,Q,R\in A^0(\mathbb{R}^3)$. 取嵌入映射 $h:N\rightarrow M$ 为恒同映射 $I:\mathbb{R}^1\rightarrow\mathbb{R}^3$,于是

$$\int_{\mathbb{R}^3}\omega = \int_{\mathbb{R}^1} h^*\circ\omega,$$

其中 $h^*\circ\omega\in A^1(\mathbb{R}^1)$在所取的局部坐标系 $u_1=x,u_2=y,u_3=z$ 之下,$h^*\circ\omega=h^*(\omega)=I^*(\omega)$,就是 $\omega=P\mathrm{d}x+Q\mathrm{d}y+R\mathrm{d}z$ 在局部坐标系上的限制,$I^*(\omega)=\omega|_U$,但由局部坐标系的取法,知

$$I^*(\omega) = I^*(P\mathrm{d}x+Q\mathrm{d}y+R\mathrm{d}z) = P\mathrm{d}x+Q\mathrm{d}y+R\mathrm{d}z,$$

故

$$h^*\circ\omega|_U = P\mathrm{d}x+Q\mathrm{d}y+R\mathrm{d}z,$$

从而

$$\int_{\mathbb{R}^1} h^*\circ\omega = \int_{\mathbb{R}^1} P\mathrm{d}u_1+Q\mathrm{d}u_2+R\mathrm{d}u_3 = \int_{\mathbb{R}^1} P\mathrm{d}x+Q\mathrm{d}y+R\mathrm{d}z.$$

例 6.4.2 是当光滑流形 $M=\mathbb{R}^3$ 的光滑子流形为 $N=\mathbb{R}^2$ 或 $N=\mathbb{R}^1$ 的例子,现在给出当 N 为低维流形,如曲面、曲线时的积分例子.

例 6.4.3 子流形上的积分.

① $M=\mathbb{R}^3$ 上的 2 次外微分式 $\omega\in A^2(\mathbb{R}^3)$在子流形 $N=S$ 为光滑曲面时的积分. 设光滑曲面的参数式为 $S:\begin{cases}x=x(u,v),\\ y=y(u,v),(u,v)\in D\subset\mathbb{R}^2,\\ z=z(u,v),\end{cases}$则曲面的法向量为

$$\vec{n} = (\cos\alpha,\cos\beta,\cos\gamma) = \frac{\left(\dfrac{\partial(y,z)}{\partial(u,v)},\dfrac{\partial(z,x)}{\partial(u,v)},\dfrac{\partial(x,y)}{\partial(u,v)}\right)}{\sqrt{\left(\dfrac{\partial(y,z)}{\partial(u,v)}\right)^2+\left(\dfrac{\partial(z,x)}{\partial(u,v)}\right)^2+\left(\dfrac{\partial(x,y)}{\partial(u,v)}\right)^2}}.$$

由 $\omega\in A^2(\mathbb{R}^3)$,$\omega=P\mathrm{d}y?\mathrm{d}z+Q\mathrm{d}z?\mathrm{d}x+R\mathrm{d}x?\mathrm{d}y$,$h:S\rightarrow\mathbb{R}^3$ 为恒同映射 $I:S\rightarrow\mathbb{R}^3$,则 $h^*\circ\omega=h^*(\omega)=I^*(\omega)$,这里

$$I^*(\omega) = I^*(P\mathrm{d}y?\mathrm{d}z+Q\mathrm{d}z?\mathrm{d}x+R\mathrm{d}x?\mathrm{d}y)$$

$$= P\left(\frac{\partial y}{\partial u}\mathrm{d}u+\frac{\partial y}{\partial v}\mathrm{d}v\right)?\left(\frac{\partial z}{\partial u}\mathrm{d}u+\frac{\partial z}{\partial v}\mathrm{d}v\right)+Q\left(\frac{\partial z}{\partial u}\mathrm{d}u+\frac{\partial z}{\partial v}\mathrm{d}v\right)?\left(\frac{\partial x}{\partial u}\mathrm{d}u+\frac{\partial x}{\partial v}\mathrm{d}v\right)$$

$$+R\left(\frac{\partial x}{\partial u}\mathrm{d}u+\frac{\partial x}{\partial v}\mathrm{d}v\right)?\left(\frac{\partial y}{\partial u}\mathrm{d}u+\frac{\partial y}{\partial v}\mathrm{d}v\right)$$

$$=P\frac{\partial(y,z)}{\partial(u,v)}\mathrm{d}u\ ?\mathrm{d}v+Q\frac{\partial(z,x)}{\partial(u,v)}\mathrm{d}u\ ?\mathrm{d}v+R\frac{\partial(x,y)}{\partial(u,v)}\mathrm{d}u\ ?\mathrm{d}v;$$

于是

$$\int_{\mathbb{R}^3}\omega=\int_S I^*\circ\omega=\int_S\left\{P\frac{\partial(y,z)}{\partial(u,v)}+Q\frac{\partial(z,x)}{\partial(u,v)}+R\frac{\partial(x,y)}{\partial(u,v)}\right\}\mathrm{d}u\ ?\mathrm{d}v$$

$$=\int_D\left\{P\frac{\partial(y,z)}{\partial(u,v)}+Q\frac{\partial(z,x)}{\partial(u,v)}+R\frac{\partial(x,y)}{\partial(u,v)}\right\}\mathrm{d}u\mathrm{d}v,$$

当面积微元记为 $\mathrm{d}S=\sqrt{\left(\frac{\partial(y,z)}{\partial(u,v)}\right)^2+\left(\frac{\partial(z,x)}{\partial(u,v)}\right)^2+\left(\frac{\partial(x,y)}{\partial(u,v)}\right)^2}\,\mathrm{d}u\mathrm{d}v$ 时，积分又记为

$$\int_{\mathbb{R}^3}\omega=\int_S\{P\cos\alpha+Q\cos\beta+R\cos\gamma\}\mathrm{d}S.$$

② $M=\mathbb{R}^3$ 上的 1 次外微分式 $\omega\in A^1(\mathbb{R}^3)$ 在子流形 $N=\gamma$ 为光滑曲线时的积分. 设光滑曲线的参数式为 $\gamma:\begin{cases}x=x(t),\\ y=y(t),\\ z=z(t),\end{cases}$ $t\in[a,b]\subset\mathbb{R}^1$，由 $\omega\in A^1(\mathbb{R}^3)$ 表示为 $\omega=P\mathrm{d}x+Q\mathrm{d}y+R\mathrm{d}z$，而 $h:\gamma\to\mathbb{R}^3$ 为恒同映射 $I:\gamma\to\mathbb{R}^3$，则 $h^*\circ\omega=h^*(\omega)=I^*(\omega)$ 化为

$$I^*(\omega)=I^*(P\mathrm{d}x+Q\mathrm{d}y+R\mathrm{d}z)=P\left(\frac{\mathrm{d}x}{\mathrm{d}t}\mathrm{d}t\right)+Q\left(\frac{\mathrm{d}y}{\mathrm{d}t}\mathrm{d}t\right)+R\left(\frac{\mathrm{d}z}{\mathrm{d}t}\mathrm{d}t\right)$$

$$=\left(P\frac{\mathrm{d}x}{\mathrm{d}t}+Q\frac{\mathrm{d}y}{\mathrm{d}t}+R\frac{\mathrm{d}z}{\mathrm{d}t}\right)\mathrm{d}t,$$

于是

$$\int_{\mathbb{R}^3}\omega=\int_\gamma I^*\circ\omega=\int_\gamma I^*(P\mathrm{d}x+Q\mathrm{d}y+R\mathrm{d}z)=\int_a^b\left(P\frac{\mathrm{d}x}{\mathrm{d}t}+Q\frac{\mathrm{d}y}{\mathrm{d}t}+R\frac{\mathrm{d}z}{\mathrm{d}t}\right)\mathrm{d}t.$$

6.4.3 Stokes 公式

在给出光滑流形的上外微分式空间中的外微分式的积分后，我们来证明流形上的 Stokes 公式.

一个区域上的积分与其边界上积分的联系，是微积分中最基本的内容之一，先回顾几个熟悉的公式.

Newton-Leibniz 公式 $\int_{[a,b]}f\mathrm{d}x=f(b)-f(a)$;

Green 公式 $\iint\limits_D\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)\mathrm{d}x\mathrm{d}y=\int_{\partial D}P\mathrm{d}x+Q\mathrm{d}y$;

Gauss 公式 $\iiint\limits_\Omega\left(\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}+\frac{\partial R}{\partial z}\right)\mathrm{d}x\mathrm{d}y\mathrm{d}z=\iint\limits_{\partial\Omega}P\mathrm{d}y\mathrm{d}z+Q\mathrm{d}z\mathrm{d}x+R\mathrm{d}x\mathrm{d}y$;

Stokes 公式

$$\iint_{\Sigma}\left(\frac{\partial R}{\partial y}-\frac{\partial Q}{\partial z}\right)\mathrm{d}y\mathrm{d}z+\left(\frac{\partial P}{\partial z}-\frac{\partial R}{\partial x}\right)\mathrm{d}z\mathrm{d}x+\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)\mathrm{d}x\mathrm{d}y=\int_{\partial\Sigma}P\mathrm{d}x+Q\mathrm{d}y+R\mathrm{d}z.$$

以上公式用外微分概念,可统一如下:

Newton-Leibniz 公式　记$[a,b]=D$,边界记为$\partial D=\{\{a\},\{b\}\}$,故公式成为

$$\int_D \mathrm{d}f=\int_{\partial D}f;$$

Green 公式　令 $\omega=P\mathrm{d}x+Q\mathrm{d}y$,则$\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)\mathrm{d}x?\mathrm{d}y=\mathrm{d}\omega$,故公式成为

$$\int_D \mathrm{d}\omega=\int_{\partial D}\omega;$$

Gauss 公式　令 $\zeta=P\mathrm{d}y\mathrm{d}z+Q\mathrm{d}z\mathrm{d}x+R\mathrm{d}x\mathrm{d}y$,则$\left(\frac{\partial P}{\partial x}+\frac{\partial Q}{\partial y}+\frac{\partial R}{\partial z}\right)\mathrm{d}x?\mathrm{d}y?\mathrm{d}z=\mathrm{d}\zeta$,故公式成为

$$\int_D \mathrm{d}\zeta=\int_{\partial D}\zeta;$$

Stokes 公式　令 $\eta=P\mathrm{d}x+Q\mathrm{d}y+R\mathrm{d}z$,则

$$\left(\frac{\partial R}{\partial y}-\frac{\partial Q}{\partial z}\right)\mathrm{d}y?\mathrm{d}z+\left(\frac{\partial P}{\partial z}-\frac{\partial R}{\partial x}\right)\mathrm{d}z?\mathrm{d}x+\left(\frac{\partial Q}{\partial x}-\frac{\partial P}{\partial y}\right)\mathrm{d}x?\mathrm{d}y=\mathrm{d}\eta,$$

故公式成为

$$\int_\Sigma \mathrm{d}\eta=\int_{\partial\Sigma}\eta.$$

为在光滑流形 M 上推广上述公式,我们再做一些准备.

1. 光滑流形 M 中的带边区域

定义 6.4.6(带边区域)　$D\subset M$ 为 m 维光滑流形 M 的子集,若 D 中的点分为两类:

(1) 内点 $p\in D$ —— 若存在点 p 的开邻域 U,使得 $p\in U\subset D$;

(2) 边界点 $s\in M$ —— 若点 s 存在一个局部坐标系$\{(U,u_j)\}$,使得 $u_j(s)=0$,并且

$$U\cap D=\{q\in U:u_m(q)\geqslant 0\},$$

则称子集 $D\subset M$ 为**带边区域(domain with boundary)**,局部坐标系$\{(U,u_j)\}$称为**边界点 $s\in M$ 的适用坐标系(adapted coordinate system)**.

区域 D 的边界点的全体,称为 D 的**边界**,记为∂D,如图 6.4.1 所示.

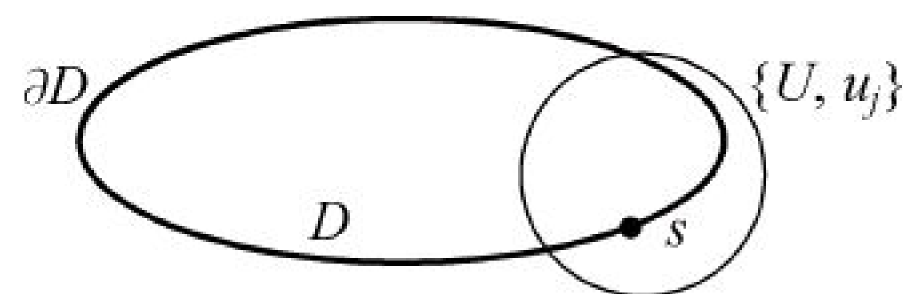


图 6.4.1　区域及其边界

注　用坐标覆盖系 $\Sigma_0=\{(U_\alpha,x_\alpha^j):\alpha\in I,1\leqslant j\leqslant m\}$表示,若一点 $p\in D$,存在 $U_\alpha\in\Sigma_0$,使 $p\in U_\alpha\subset D$,则称 p 为**内点**;若一点 $s\in M$,存在 $U_\alpha\in\Sigma_0$,使$\forall\,1\leqslant j\leqslant m$,都有 $x_\alpha^j(\varphi(s))=0$ 成

立，其中$(U_\alpha,\varphi_\alpha)$为 s 点的坐标卡，并且

$$\varphi(U\cap D)=\varphi(U)\cap\{(x^1,\cdots,x^m)\in\mathbb{R}^m:\ x^m\geqslant 0\}=\varphi(U)\cap\mathbb{R}^m,$$

则称 s 点为 D 的**边界点**；称 D 为 M 的**带边区域**. 于是，有

$$\varphi(U\cap\partial D)=\varphi(U)\cap\{(x^1,\cdots,x^m)\in\mathbb{R}^m:x^m=0\}.$$

定理 6.4.4（带边区域边界的定向定理）　设 D 为光滑流形 M 的带边区域，则其边界 ∂D 是 M 的 $m-1$ 维正则嵌入闭子流形. 若 M 是可定向流形，则∂D 也是可定向的.

证　显然，D 的边界∂D 是 M 的闭子集. 设$\{(U,u_j)\}$是边界点 $s\in M$ 的适用坐标系，则 $U\cap\partial D=\{q\in U:u_m(q)=0\}$，根据定义，$\partial D$ 是 M 的正则嵌入的闭子流形.

由 M 的定向来确定∂D 的定向 —— 对于任一点 $s\in\partial D$，选取与 M 的定向相符合的适用坐标系$\{(U,u_j)\}$，使得 $\mathrm{d}u_1$?⋯?$\mathrm{d}u_{m-1}$?$\mathrm{d}u_m$ 是 M 的定向. 则$(u_1,\cdots,u_{m-1})$是边界∂D 在点 $s\in\partial D$ 的局部坐标. 于是，可以用

$$(-1)^m\mathrm{d}u_1?\cdots?\mathrm{d}u_{m-1}\tag{6.4.6}$$

给出边界∂D（$m-1$ 维）在点 $s\in\partial D$ 的坐标区域 $U\cap\partial D=\{q\in U:u_m(q)=0\}$上的定向.

我们证明，如上给出的坐标系的定向是彼此相容的.

设$\{(W,w_k)\}$是边界点 $s\in\partial D$ 的另一组与 M 的定向相适用的坐标系，则

$$\frac{\partial(w_1,\cdots,w_{m-1},w_m)}{\partial(u_1,\cdots,u_{m-1},u_m)}>0.\tag{6.4.7}$$

在不同的坐标系中，若设 $w_m=f_m(u_1,\cdots,u_{m-1},u_m)$，则对于任意固定的 $u_1,\cdots,u_{m-1}$，变量 w_m 的符号与 u_m 的符号相同，并且 $u_m=0\Rightarrow w_m=0$. 所以，在点 $s\in\partial D$ 有 $\dfrac{\partial w_m}{\partial u_m}>0$. 于是，不失一般性，可设 $w_m=u_m$. 这样就可从(6.4.7)式得到

$$\frac{\partial(w_1,\cdots,w_{m-1})}{\partial(u_1,\cdots,u_{m-1})}>0.\tag{6.4.8}$$

这说明，相应的$(-1)^m\mathrm{d}u_1$?⋯?$\mathrm{d}u_{m-1}$与$(-1)^m\mathrm{d}w_1$?⋯?$\mathrm{d}w_{m-1}$在 $U\cap W\cap\partial D$ 上给出的定向是一致的. 所以，∂D 是可以定向的.

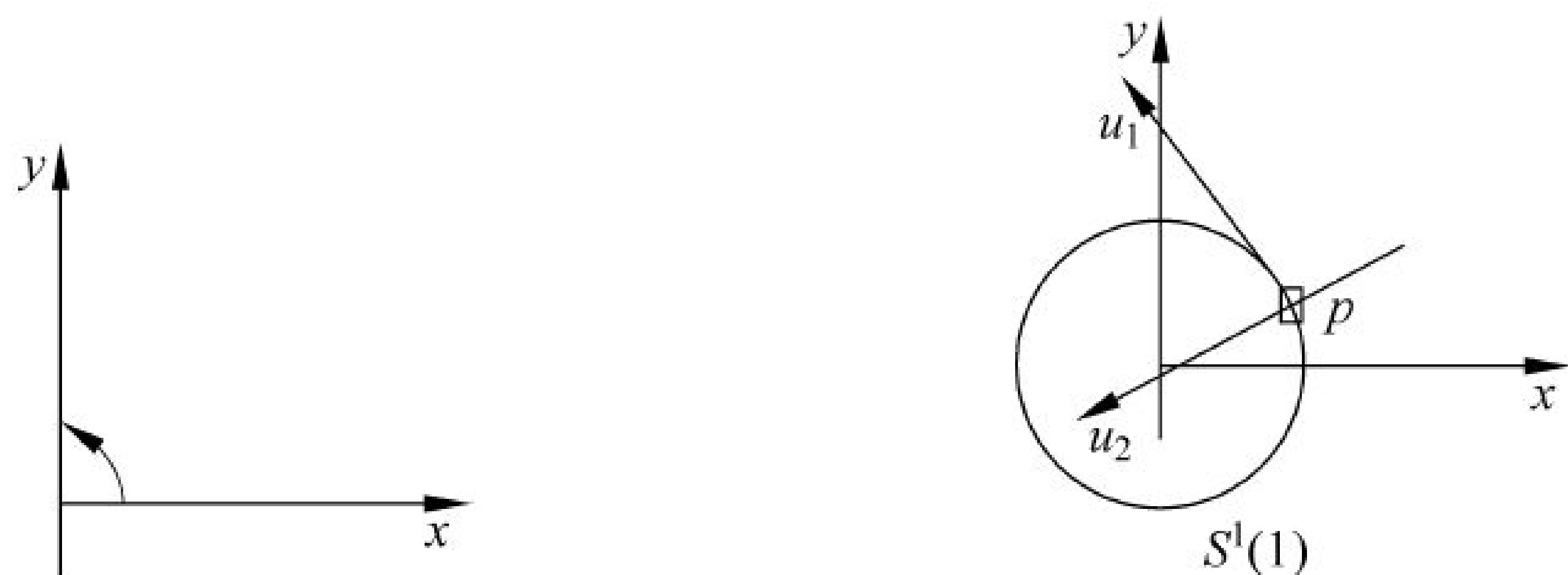
定义 6.4.7（边界的定向，诱导定向）　对于 m 维定向光滑流形 M 的带边区域 $D\subset M$，其边界∂D 的方向由$(-1)^m\mathrm{d}u_1$?⋯?$\mathrm{d}u_{m-1}$确定，其中$\{(U,u_j)\}$是 $s\in\partial D$ 的适用坐标系，$(u_1,\cdots,u_{m-1})$是边界∂D 在点 $s\in\partial D$ 的局部坐标系. 由(6.4.6)式给出的定向$(-1)^m\mathrm{d}u_1$?⋯?$\mathrm{d}u_{m-1}$称为**定向光滑流形 M 在∂D 上的诱导定向**.

通常，把流形 M 的定向视为带边区域 D 的定向，于是，带边区域 D 及其边界∂D 的诱导定向便完全确定.

回顾在 $\mathbb{R}^3$ 的情形，平面区域 D 及其边界∂D、空间区域 Ω 及其边界$\partial\Omega$、空间曲面 Σ 及其边界$\partial\Sigma$（是空间曲线）正是定义 6.4.7 的特例.

例 6.4.4

① 单位闭圆盘 $B_0(1)=\{(x,y)\in\mathbb{R}^2: x^2+y^2\leqslant 1\}$是 $\mathbb{R}^2$ 中的带边区域,边界为 $\mathbb{R}^2$ 中的单位圆周 $S^1(1)=\{(x,y)\in\mathbb{R}^2: x^2+y^2=1\}=\partial B_0(1)$. 取 $M=\mathbb{R}^2$,$\omega\in A^2(\mathbb{R}^2)\Rightarrow\omega=a\mathrm{d}x?\mathrm{d}y$,$a\in C^\infty(\mathbb{R}^3)\Rightarrow$ 取 $\omega=\mathrm{d}x?\mathrm{d}y$(右手系,即逆时针向). 取 $B_0(1)$的适用局部坐标系为$\{(U,(u_1,u_2))\}$,其定向为 $\mathrm{d}u_1?\mathrm{d}u_2$. 按定义 6.4.7,在$\partial B_0(1)$上的诱导定向为$(-1)^2\mathrm{d}u_1=\mathrm{d}u_1$.



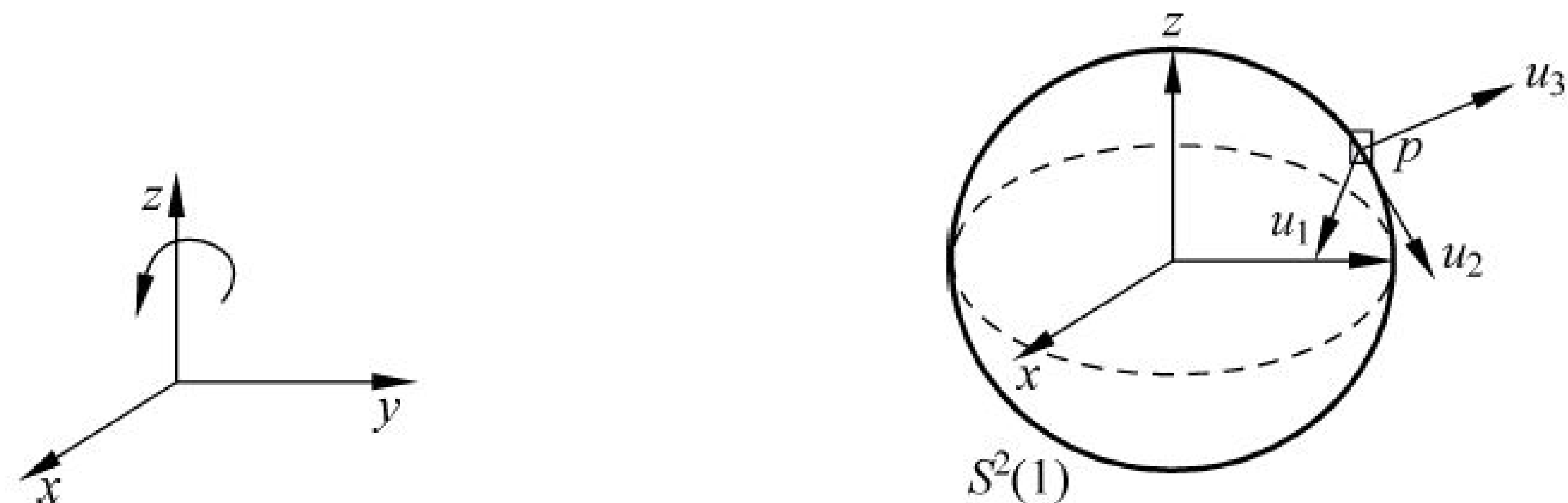
以上定义的$\partial B_0(1)$上的定向与我们通常使用的$\partial B_0(1)$上的定向是一致的. 因为通常$\partial B_0(1)$的定向规定为:当行进者沿边界$\partial B_0(1)$正向(逆时针向)行进时,则区域 $B_0(1)$本身在行进者的左侧. 这个侧与适用局部坐标系$\{(U,(u_1,u_2))\}$的关系为:$\frac{\partial}{\partial u_1}$与$\partial B_0(1)$在点 p 相切,因此$\frac{\partial}{\partial u_1}$是点 p 的切标架;而$\frac{\partial}{\partial u_2}$指向 $B_0(1)$的内部,亦即$\frac{\partial}{\partial u_2}$指向沿$\partial B_0(1)$的正向行进者的左侧. 因此,由$(-1)^2\mathrm{d}u_1=\mathrm{d}u_1$ 作为$\partial B_0(1)$的定向,这与通常使用的定向一致.

② 对于 $M=\mathbb{R}^3$, $\omega\in A^3(\mathbb{R}^3)\Rightarrow\omega=a\mathrm{d}x?\mathrm{d}y?\mathrm{d}z$,$a\in C^\infty(\mathbb{R}^3)$,取 $\omega=1\cdot\mathrm{d}x?\mathrm{d}y?\mathrm{d}z=\mathrm{d}x?\mathrm{d}y?\mathrm{d}z$ 为$\mathbb{R}^3$ 的定向.

$\mathbb{R}^3$ 中的单位闭球 $B_0(1)=\{(x,y,z)\in\mathbb{R}^3: x^2+y^2+z^2\leqslant 1\}$的边界为单位球面

$$\partial B_0(1)\equiv S^2(1)=\{(x,y,z)\in\mathbb{R}^3: x^2+y^2+z^2=1\}.$$

当$\mathbb{R}^3$ 的定向确定后,$\omega=\mathrm{d}x?\mathrm{d}y?\mathrm{d}z$(右手系),通常取点 $p\in\partial B_0(1)$的右手单位正交容许局部坐标系$\{(U_p,(e_1,e_2,e_3))\}$,使得(e_1,e_2,e_3)决定的方向与 $\omega=\mathrm{d}x?\mathrm{d}y?\mathrm{d}z$ 右手系一致;亦即,e_1,e_2 是边界$\partial B_0(1)$的切矢量,而 e_3 是$\partial B_0(1)$的指向区域 $B_0(1)$的外法线方向. 于是,$\partial B_0(1)$在点 $p\in\partial B_0(1)$的切空间上的定向用切标架$(p;e_1,e_2)$给出,就是用点 $p\in\partial B_0(1)$的外法线方向作为$\partial B_0(1)$的定向.



现在用流形、子流形的方法来确定$\partial B_0(1)$的方向. 取$B_0(1)$的适用局部坐标系$\{(U_p,(u_1,u_2,u_3))\}$,其定向$\mathrm{d}u_1?\mathrm{d}u_2?\mathrm{d}u_3$与$\omega=\mathrm{d}x?\mathrm{d}y?\mathrm{d}z$一致. 在$\partial B_0(1)$上的定向由$(-1)^3\mathrm{d}u_1?\mathrm{d}u_2=-\mathrm{d}u_1?\mathrm{d}u_2$给出. 这个方向与通常所给的外法向一致,这是因为:由适用局部坐标系$\{(U_p,(u_1,u_2,u_3))\}$生成余切空间的局部坐标系为$\{(p,(\mathrm{d}u_1)_p,(\mathrm{d}u_2)_p,(\mathrm{d}u_3)_p)\}$,切空间的局部坐标系为$\left\{\left(p,\left.\frac{\partial}{\partial u_1}\right|_p,\left.\frac{\partial}{\partial u_2}\right|_p,\left.\frac{\partial}{\partial u_3}\right|_p\right)\right\}$,其中$\left\{p,\left.\frac{\partial}{\partial u_1}\right|_p,\left.\frac{\partial}{\partial u_2}\right|_p\right\}$构成$\partial B_0(1)$的切标架,$\left.\frac{\partial}{\partial u_3}\right|_p$与$\partial B_0(1)$横截,并且指向区域$B_0(1)$的内部. 如将标架中的$\left.\frac{\partial}{\partial u_1}\right|_p$,$\left.\frac{\partial}{\partial u_3}\right|_p$同时反向,则标架$\left\{p,\left.\frac{\partial}{\partial u_1}\right|_p,\left.\frac{\partial}{\partial u_2}\right|_p,\left.\frac{\partial}{\partial u_3}\right|_p\right\}$与$\left\{p,-\left.\frac{\partial}{\partial u_1}\right|_p,\left.\frac{\partial}{\partial u_2}\right|_p,-\left.\frac{\partial}{\partial u_3}\right|_p\right\}$仍与$\mathbb{R}^3$的原定向$\omega=\mathrm{d}x?\mathrm{d}y?\mathrm{d}z$保持一致,但是,$\left.\frac{\partial}{\partial u_3}\right|_p$虽仍与$\partial B_0(1)$横截,却指向区域$B_0(1)$的外部. 所以,在切空间$T_p$上,由外微分式$-\mathrm{d}u_1?\mathrm{d}u_2$给出的边界的诱导定向恰好是由切标架$\left\{p,-\left.\frac{\partial}{\partial u_1}\right|_p,\left.\frac{\partial}{\partial u_2}\right|_p\right\}$给出的定向,也就是“$\partial B_0(1)$的外法向为正定向”.

2. 流形上的Stokes公式

定理6.4.5 设M为m维定向光滑流形,$D\subset M$是M中的带边区域,边界∂D取诱导定向;ω是M上的具有紧致支集的$m-1$次外微分式,则

$$\int_D \mathrm{d}\omega=\int_{\partial D}\omega. \tag{6.4.9}$$

证 设$\{\{U_j\},\{g_\alpha\}\}$是与M的定向相符的坐标覆盖与从属的单位分解,则$\omega=\sum\limits_\alpha g_\alpha\cdot\omega$. 因为$\operatorname{supp}\omega$是紧致集,等式$\omega=\sum\limits_\alpha g_\alpha\cdot\omega$右边只是有限和. 于是$\int_D\mathrm{d}\omega=\sum\limits_\alpha\int_D\mathrm{d}(g_\alpha\cdot\omega)$,且$\int_{\partial D}\omega=\sum\limits_\alpha\int_{\partial D}g_\alpha\cdot\omega$. 这说明,只要对于每一个$\alpha$,证明等式$\int_D\mathrm{d}(g_\alpha\cdot\omega)=\int_{\partial D}g_\alpha\cdot\omega$即可.

不失一般性,设支集$\operatorname{supp}\omega$包含在$M$的一个与定向相符的局部坐标系$\{(U,u_j)\}$内. 设$\omega$的表示式为

$$\omega=\sum_{j=1}^m(-1)^j a_j\,\mathrm{d}u_1?\cdots?\mathrm{d}\hat{u}_j?\cdots?\mathrm{d}u_m,$$

其中$a_j\in A^0(M)$,$\mathrm{d}u_1?\cdots?\mathrm{d}\hat{u}_j?\cdots?\mathrm{d}u_m\equiv\mathrm{d}u_1?\cdots?\mathrm{d}u_{j-1}?\mathrm{d}u_{j+1}?\cdots?\mathrm{d}u_m$,则

$$\mathrm{d}\omega=\left(\sum_{j=1}^m\frac{\partial a_j}{\partial u_j}\right)\mathrm{d}u_1?\cdots?\mathrm{d}u_m.$$

下面分两种情形讨论:

(1) 若$U\cap\partial D=\varnothing$,则(6.4.9)式的右边为零. 考察左边$\int_D\mathrm{d}\omega$. 当$U\subset M\backslash D$时,显然$\int_D\mathrm{d}\omega=0$;当$U\subset D$时,有

$$\int_{D} \mathrm{d}\omega = \int_{D} \mathrm{d}\omega = \int_{D} \left(\sum_{j=1}^{m} \frac{\partial a_j}{\partial u_j} \right) \mathrm{d}u_1 ? \cdots ? \mathrm{d}u_m .$$

考虑$\mathbb{R}^m$ 中的一个方体 $K=\{|u_j|\leqslant c, 1\leqslant j\leqslant m\}$，使得 $U\subset K$. 将函数 $a_j, 1\leqslant j\leqslant m$，延拓到 K 上，使得在 $K\backslash U$ 上为零，并保持 a_j 在 K 内连续可微，因此有

$$\int_U \frac{\partial a_j}{\partial u_j}\mathrm{d}u_1\cdots\mathrm{d}u_m = \int_K \frac{\partial a_j}{\partial u_j}\mathrm{d}u_1\cdots\mathrm{d}u_m = \int_U \frac{\partial a_j}{\partial u_j}\mathrm{d}u_1\cdots\mathrm{d}u_m$$

$$= \int_{|u_j|\leqslant c} \left(\int_{-c}^{c} \frac{\partial a_j}{\partial u_j}\mathrm{d}u_j \right) \mathrm{d}u_1\cdots\mathrm{d}u_{j-1}\mathrm{d}u_{j+1}\cdots\mathrm{d}u_m = 0,$$

最后一项为零，是因为

$$\int_{-c}^{c} \frac{\partial a_j}{\partial u_j}\mathrm{d}u_j = a_j(u_1,\cdots,u_{j-1},c,u_{j+1},\cdots,u_m) - a_j(u_1,\cdots,u_{j-1},-c,u_{j+1},\cdots,u_m) = 0.$$

(2) 若 $U\cap\partial D\neq\varnothing$，设 U 就是与 M 的定向相符的适用坐标域，即 $U\cap D=\{q\in U: u_m(q)\geqslant 0\}$，且 $U\cap\partial D=\{q\in U: u_m(q)=0\}$. 类似地，在 $\mathbb{R}^m$ 中取一个方体

$$K=\{|u_j|\leqslant c, 1\leqslant j\leqslant m-1; 0\leqslant u_m\leqslant c\},$$

于是，当常数 $c>0$ 充分大时，$U\cap D$ 将落在 K 的内部与边界 $u_m=0$ 的并集中. 如同(1)一样，将函数 a_j 延拓到 K 上，使得在 $U\cap D$ 之外为零，并保持 a_j 在 K 内连续可微，因此，(6.4.9)式右边为

$$\int_{\partial D}\omega = \int_{U\cap\partial D}\omega = \sum_{j=1}^{m}(-1)^{j-1}\int_{U\cap\partial D} a_j \mathrm{d}u_1 ?\cdots ?\mathrm{d}u_{j-1}?\mathrm{d}u_{j+1}?\cdots?\mathrm{d}u_m$$

$$= (-1)^{m-1}\int_{U\cap\partial D} a_m \mathrm{d}u_1 ?\cdots?\mathrm{d}u_{m-1} \text{（在 } U\cap\partial D \text{ 上 } \mathrm{d}u_m=0\text{）}$$

$$= -\int_{|u_j|\leqslant c, 1\leqslant j\leqslant m-1} a_m(u_1,\cdots,u_{m-1},0)\mathrm{d}u_1\cdots\mathrm{d}u_{m-1}. \text{（由 } M \text{ 在 } \partial D \text{ 上的诱导定向）}$$

这样，(6.4.9)式左边为

$$\int_D \mathrm{d}\omega = \int_{D\cap U}\mathrm{d}\omega = \sum_{j=1}^{m}\int_{D\cap U}\frac{\partial a_j}{\partial u_j}\mathrm{d}u_1 ?\cdots?\mathrm{d}u_m. \tag{6.4.10}$$

对于 $1\leqslant j\leqslant m-1$，则

$$\int_{D\cap U}\frac{\partial a_j}{\partial u_j}\mathrm{d}u_1 ?\cdots?\mathrm{d}u_m = \int_{|u_j|\leqslant c, j\neq m, 0\leqslant u_m\leqslant c}\left(\int_{-c}^{c}\frac{\partial a_j}{\partial u_j}\mathrm{d}u_j\right)\mathrm{d}u_1\cdots\mathrm{d}u_{j-1}\mathrm{d}u_{j+1}\cdots\mathrm{d}u_m = 0,$$

所以，(6.4.10)式右边只有一项不为零，即

$$\int_{D\cap U}\frac{\partial a_m}{\partial u_m}\mathrm{d}u_1 ?\cdots?\mathrm{d}u_m = \int_{|u_j|\leqslant c, j\neq m}(a_m(u_1,\cdots,u_{m-1},c) - a_m(u_1,\cdots,u_{m-1},0))\mathrm{d}u_1\cdots\mathrm{d}u_{m-1}$$

$$= -\int_{|u_j|\leqslant c, j\neq m} a_m(u_1,\cdots,u_{m-1},0)\mathrm{d}u_1\cdots\mathrm{d}u_{m-1}.$$

因此左右相等，定理得证.

在实际应用中，经常处理的情况是带边区域 D 是闭区域，而且是紧致集，所以，不必假

设 $m-1$ 次外微分式具有紧致支集，Stokes 公式仍然成立. Stokes 公式在物理、力学、偏微分方程、微分几何学中都有十分重要的应用.

定理 6.4.2 的注中的从属单位分解起着重要作用，扮演不可缺少的角色：一方面是“分解”，把问题化到每个局部坐标系上处理，另一方面是“综合”，把各个坐标系上已经做好的表示式综合起来，成为定义在整个光滑流形上的数学对象.

例 6.4.5 当 $S=S^2(1)=\{(x,y,z)\in\mathbb{R}^3: x^2+y^2+z^2=1\}$ 时，直接计算、用 Stokes 公式计算积分 $\int_{\mathbb{R}^3}\omega=\int_{S^2(1)}I^*\circ\omega$，其中 $\omega=x\mathrm{d}y?\mathrm{d}z+y\mathrm{d}z?\mathrm{d}x+z\mathrm{d}x?\mathrm{d}y$.

① 直接计算. 在 $S^2(1)$ 上取局部坐标系

$$U=\{(x,y,z)\in S^2(1): z\geqslant 0\},\quad \{U,(u_1,u_2)\},$$
$$V=\{(x,y,z)\in S^2(1): z\leqslant 0\},\quad \{V,(v_1,v_2)\},$$

令

$$\varphi_U(u_1,u_2)=(u_1,u_2,\sqrt{1-u_1^2-u_2^2}),\quad \varphi_U: U\to\varphi_U(U),\ \varphi_U(U)\subset\mathbb{R}^3,$$
$$\varphi_V(v_1,v_2)=(v_1,v_2,-\sqrt{1-v_1^2-v_2^2}),\quad \varphi_V: V\to\varphi_V(V),\ \varphi_V(V)\subset\mathbb{R}^3.$$

则 $S^2(1)\subset U\cup V$，于是

$$\begin{aligned}
&I^*(x\mathrm{d}y?\mathrm{d}z+y\mathrm{d}z?\mathrm{d}x+z\mathrm{d}x?\mathrm{d}y)\big|_U\\
={}&u_1\mathrm{d}u_2?\left(\frac{\partial z}{\partial u_1}\mathrm{d}u_1+\frac{\partial z}{\partial u_2}\mathrm{d}u_2\right)+u_2\left(\frac{\partial z}{\partial u_1}\mathrm{d}u_1+\frac{\partial z}{\partial u_2}\mathrm{d}u_2\right)?\mathrm{d}u_1\\
&+\sqrt{1-u_1^2-u_2^2}\,\mathrm{d}u_1?\mathrm{d}u_2\\
={}&\mathrm{d}u_2?\left(\frac{-u_1^2}{\sqrt{1-u_1^2-u_2^2}}\mathrm{d}u_1+\frac{-u_1u_2}{\sqrt{1-u_1^2-u_2^2}}\mathrm{d}u_2\right)\\
&+\left(\frac{-u_2u_1}{\sqrt{1-u_1^2-u_2^2}}\mathrm{d}u_1+\frac{-u_2^2}{\sqrt{1-u_1^2-u_2^2}}\mathrm{d}u_2\right)?\mathrm{d}u_1\\
&+\sqrt{1-u_1^2-u_2^2}\,\mathrm{d}u_1?\mathrm{d}u_2\\
={}&\frac{u_1^2}{\sqrt{1-u_1^2-u_2^2}}\mathrm{d}u_1?\mathrm{d}u_2+\frac{u_2^2}{\sqrt{1-u_1^2-u_2^2}}\mathrm{d}u_2?\mathrm{d}u_1\\
&+\frac{1-u_1^2-u_2^2}{\sqrt{1-u_1^2-u_2^2}}\mathrm{d}u_1?\mathrm{d}u_2,
\end{aligned}$$

故

$$I^*(x\mathrm{d}y?\mathrm{d}z+y\mathrm{d}z?\mathrm{d}x+z\mathrm{d}x?\mathrm{d}y)\big|_U=\frac{1}{\sqrt{1-u_1^2-u_2^2}}\mathrm{d}u_1?\mathrm{d}u_2,$$

同理有

$$I^*(x\mathrm{d}y?\mathrm{d}z+y\mathrm{d}z?\mathrm{d}x+z\mathrm{d}x?\mathrm{d}y)\big|_V=\frac{1}{\sqrt{1-v_1^2-v_2^2}}\mathrm{d}v_1?\mathrm{d}v_2.$$

注意到 $U\cap V=\{(x,y,z)\in\mathbb{R}^3: z=0\}$ 是赤道，故其 Lebesgue 测度为 0，于是

$$\int_{S^2(1)} I^* \circ \omega = \int_U I^* \circ \omega + \int_V I^* \circ \omega,$$

对于 U,使用极坐标系

$$\begin{cases} u_1 = r\cos\theta, \\ u_2 = r\sin\theta, \end{cases} \quad 0 \leqslant r \leqslant 1, -\pi \leqslant \theta \leqslant \pi,$$

因此

$$\mathrm{d}u_1 ? \mathrm{d}u_2 = \left(\frac{\partial u_1}{\partial r}\mathrm{d}r + \frac{\partial u_1}{\partial \theta}\mathrm{d}\theta\right) ? \left(\frac{\partial u_2}{\partial r}\mathrm{d}r + \frac{\partial u_2}{\partial \theta}\mathrm{d}\theta\right) = r\,\mathrm{d}r ? \mathrm{d}\theta,$$

但 $U \leftrightarrow D_0 = D \backslash \{(u,0) \in D : -1 \leqslant u \leqslant 0\}$,因此

$$\begin{aligned} \int_U I^* \circ \omega &= \int_U \frac{1}{\sqrt{1-u_1^2-u_2^2}}\mathrm{d}u_1 ? \mathrm{d}u_2 = \int_{D_0} \frac{r}{\sqrt{1-r^2}}\mathrm{d}r ? \mathrm{d}\theta \\ &= \int_{-\pi}^{\pi}\int_0^1 \frac{r}{\sqrt{1-r^2}}\mathrm{d}r\mathrm{d}\theta = 2\pi, \end{aligned}$$

最后有

$$\int_{S^2(1)} I^* \circ \omega = \int_U I^* \circ \omega + \int_V I^* \circ \omega = 4\pi.$$

② 用 Stokes 公式计算. $S = S^2(1) = \{(x,y,z) \in \mathbb{R}^3 : x^2+y^2+z^2=1\}$所围的区域为

$$B_0(1) = \{(x,y,z) \in \mathbb{R}^3 : x^2+y^2+z^2 \leqslant 1\},$$

于是,积分$\int_{S^2(1)} I^* \circ \omega = \int_{B_0(1)} \mathrm{d}\omega$.

由 $\omega = x\mathrm{d}y ? \mathrm{d}z + y\mathrm{d}z ? \mathrm{d}x + z\mathrm{d}x ? \mathrm{d}y$ 得

$$\begin{aligned} \mathrm{d}\omega &= \mathrm{d}(x\mathrm{d}y ? \mathrm{d}z + y\mathrm{d}z ? \mathrm{d}x + z\mathrm{d}x ? \mathrm{d}y) \\ &= \mathrm{d}x ? \mathrm{d}y ? \mathrm{d}z + \mathrm{d}y ? \mathrm{d}z ? \mathrm{d}x + \mathrm{d}z ? \mathrm{d}x ? \mathrm{d}y \\ &= 3\mathrm{d}x ? \mathrm{d}y ? \mathrm{d}z, \end{aligned}$$

积分成为

$$\int_{S^2(1)} I^* \circ \omega = \int_{B_0(1)} \mathrm{d}\omega = \int_{B_0(1)} 3\mathrm{d}x ? \mathrm{d}y ? \mathrm{d}z = 3\int_{B_0(1)} \mathrm{d}x ? \mathrm{d}y ? \mathrm{d}z,$$

取 $B_0(1)$的球坐标系

$$\begin{cases} x = r\cos\varphi\cos\theta, \\ y = r\cos\varphi\sin\theta, \\ z = r\sin\varphi, \end{cases} \quad 0 \leqslant r \leqslant 1, 0 \leqslant \theta \leqslant 2\pi, -\frac{\pi}{2} \leqslant \varphi \leqslant \frac{\pi}{2},$$

则

$$\begin{aligned} \mathrm{d}x ? \mathrm{d}y ? \mathrm{d}z &= \begin{vmatrix} \cos\varphi\cos\theta & -r\cos\varphi\sin\theta & -r\sin\varphi\cos\theta \\ \cos\varphi\sin\theta & r\cos\varphi\cos\theta & -r\sin\varphi\sin\theta \\ \sin\varphi & 0 & r\cos\varphi \end{vmatrix} \mathrm{d}r ? \mathrm{d}\theta ? \mathrm{d}\varphi \\ &= r^2\cos\varphi\mathrm{d}r ? \mathrm{d}\theta ? \mathrm{d}\varphi. \end{aligned}$$

计算积分

$$\int_{S^2(1)} I^* \circ \omega = \int_{B_0(1)} \mathrm{d}\omega = 3\int_{B_0(1)} \mathrm{d}x ? \mathrm{d}y ? \mathrm{d}z = 3\int_{B_0(1)} r^2 \cos\varphi \mathrm{d}r ? \mathrm{d}\theta ? \mathrm{d}\varphi$$

$$= 3\int_{B_0(1)} r^2 \cos\varphi \mathrm{d}r \mathrm{d}\theta \mathrm{d}\varphi = 3\int_0^1 r^2 \mathrm{d}r \int_0^{2\pi} \mathrm{d}\theta \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos\varphi \mathrm{d}\varphi = 3 \times \frac{4\pi}{3} = 4\pi.$$

注 直接用

$$\mathrm{d}x ? \mathrm{d}y ? \mathrm{d}z = \left(\frac{\partial x}{\partial r}\mathrm{d}r + \frac{\partial x}{\partial \theta}\mathrm{d}\theta + \frac{\partial x}{\partial \varphi}\mathrm{d}\varphi\right) ? \left(\frac{\partial y}{\partial r}\mathrm{d}r + \frac{\partial y}{\partial \theta}\mathrm{d}\theta + \frac{\partial y}{\partial \varphi}\mathrm{d}\varphi\right) ? \left(\frac{\partial z}{\partial r}\mathrm{d}r + \frac{\partial z}{\partial \theta}\mathrm{d}\theta + \frac{\partial z}{\partial \varphi}\mathrm{d}\varphi\right)$$

得到同样结果. 此时,极坐标系$\left\{\frac{\partial}{\partial r}, \frac{\partial}{\partial \theta}, \frac{\partial}{\partial \varphi}\right\}$成右手系,此定向与$\mathbb{R}^3$的定向 $\mathrm{d}x ? \mathrm{d}y ? \mathrm{d}z$ 一致.

3. 光滑流形 M 上的积分算子

设$(M, \mathbb{A})$为满足第二可数公理的局部紧致的定向 m 维光滑流形,积分算子

$$\int_M : A^m(M) \to \mathbb{R}$$

是 $A^m(M)$上的线性泛函,它有如下性质.

定理 6.4.6 积分算子$\int_M : A^m(M) \to \mathbb{R}$ 是线性泛函,即

(1) $\int_M \omega^1 + \omega^2 = \int_M \omega^1 + \int_M \omega^2$, $\omega^1, \omega^2 \in A^m(M)$;

(2) $\int_M \alpha\omega = \alpha \int_M \omega$, $\omega \in A^m(M)$, $\alpha \in \mathbb{R}$.

6.5 Riemann 流形、数学科学与现代物理

6.5.1 Riemann 流形

现在我们来建立 Riemann 流形.

1. 非退化张量场,正定张量场 G

设 $M \equiv (M, \mathbb{A})$是 m 维光滑流形,G 是 M 上对称的二阶协变张量场. 若$\{(U, u_j)\}$是 M 的一个局部坐标系,则 G 可表示为

$$G = g_{ij} \mathrm{d}u_i \otimes \mathrm{d}u_j, \tag{6.5.1}$$

其中 $g_{ij} = g_{ji}$是 U 上的光滑函数. 于是,G 在 M 上的每一点 $p \in M$ 给出了 $T_p(M)$上的二重线性函数 $G(X, Y)$. 设 $X = X_i \frac{\partial}{\partial u_i}, Y = Y_j \frac{\partial}{\partial u_j}$,则

$$G(X, Y) = g_{ij} X_i Y_j. \tag{6.5.2}$$

和式约定 上式左边不出现 i,j,表示右边已经对 $i=1,2,\cdots,m,j=1,2,\cdots,m$ 求和,以下都用此约定.

称张量场 G 在点 $p\in M$ 为**非退化的**(**non-degenerated**),若切矢量 $X\in T_p(M)$ 满足:$\forall Y\in T_p(M)$ 成立 $G(X,Y)=0$ 时,必有 $X=0$.

命题 6.5.1(**非退化张量场的充分必要条件**) 张量场 G 在点 p 为非退化的,当且仅当方程组

$$g_{ij}(p)X_i=0,\quad 1\leqslant j\leqslant m$$

只有零解,也有,当且仅当行列式 $\det[g_{ij}(p)]\neq 0$.

定义 6.5.1(**正定张量场**) 对于(6.5.2)式中的 $G(X,Y)=g_{ij}X_iX_j$,若 $\forall X\in T_p(M)$,都有 $G(X,X)\geqslant 0$,且等号成立当且仅当 $X=0$,则称张量场 G 在点 $p\in M$ 是**正定的**(**positive definite**).

由线性代数知识可知,G 是正定的,当且仅当方阵$[g_{ij}]$是正定的. 因此正定张量场必定是非退化的.

2. 广义 Riemann 流形,Riemann 流形

定义 6.5.2(**广义 Riemann 流形,Riemann 流形**) 若在 m 维光滑流形 M 上给定一个光滑的、处处非退化的、对称的、二阶协变张量场 G,则称 M **是广义 Riemann 流形**(**generalized Riemann manifold**);G 称为该广义 Riemann 流形 M 的**基本张量场**,或**度量张量场**.

若 G 是光滑的、处处非退化的、对称的、正定的、二阶协变张量场,则称 M 为 **Riemann 流形**,G 仍称为 Riemann 流形 M 的基本张量场,或度量张量场.

记广义 Riemann 流形或 Riemann 流形为(M,G).

大家记得,在前几节中介绍流形时,给出了一个集合 M 的拓扑结构、微分结构,使得 M 成为一个微分流形,但是却没有给出它的度量. 现在,可以对于定义 6.5.2 中的 Riemann 流形(M,G),给出其上的度量.

由于(6.5.2)式,可以在每一点 $p\in M$ 的切空间 $T_p(M)$中定义内积如下.

对于 $X,Y\in T_p(M)$,令 $X\cdot Y=G(X,Y)=g_{ij}(p)X_iY_j$. 当 G 是正定的情形下,切矢量 $X\in T_p(M)$的**长度**可定义为

$$\|X\|=\sqrt{X\cdot X};\tag{6.5.3}$$

两个切矢量 $X,Y\in T_p(M)$在同一点 $p\in M$ 的夹角$\angle(X,Y)$的余弦可定义为

$$\cos\angle(X,Y)=\frac{X\cdot Y}{\|X\|\|Y\|},\tag{6.5.4}$$

其中 $0\leqslant\angle(X,Y)\leqslant\pi$.

于是,**Riemann 流形**(M,G)**是在** M **的每一点的切空间** $T_p(M)$**上给定了一个正定内积** $X\cdot Y$ **的微分流形**,且当 X,Y 是光滑切矢量场时,内积 $X\cdot Y$ 就是 M 上的光滑函数.

3. Riemann 度量,弧长微元

定义 6.5.3(Riemann 度量、弧长微元) 称二次微分式

$$ds^2 = g_{ij}\,du_i du_j \equiv g_{ij}\,du_i \otimes du_j \tag{6.5.5}$$

为 **Riemann 度量(Riemann metric)**,或度量形式.

$ds^2 = g_{ij}\,du_i du_j$ 与局部坐标系$\{(U,u_i)\}$的选取无关. $ds = \sqrt{g_{ij}\,du_i du_j}$ 恰好是无穷小切矢量的长度,称为**弧长微元(infinitesimal of arc)**.

设 $\gamma: u_i = u_i(t), t_0 \leqslant t \leqslant T$ 是光滑流形 M 上的一条连续的分段光滑的参数曲线,则 γ 的弧长为

$$s = \int_{t_0}^{T} \sqrt{g_{ij} \frac{du_i}{dt} \frac{du_j}{dt}}\,dt. \tag{6.5.6}$$

定理 6.5.1(Riemann 度量的存在性定理) 在 m 维光滑流形$(M,\mathbb{A})$上必定存在 Riemann 度量.

证 取光滑流形 M 的局部有限坐标覆盖$\{(U_\alpha, u_i^\alpha)\}$,设$\{h_\alpha\}$是相应的单位分解,使支集 $\mathrm{supp}\,h_\alpha \subset U_\alpha$. 令 $ds_\alpha^2 = \sum_{i=1}^{m} (du_i^\alpha)^2$,记

$$ds^2 = \sum_\alpha h_\alpha \cdot ds_\alpha^2, \tag{6.5.7}$$

其中 $h_\alpha \cdot ds_\alpha^2$ 为 $(h_\alpha \cdot ds_\alpha^2)(p) = \begin{cases} h_\alpha(p) \cdot ds_\alpha^2, & p \in U, \\ 0, & p \notin U, \end{cases}$ 它们是 M 上的光滑二次微分式.

对于每一点 $p \in M$,(6.5.7)式中的求和仅是有限项和,因为实际上总可取 p 的一个局部坐标系$\{(U,u_j)\}$,使得 $\overline{U}$ 是紧致的. 又由于$\{U_\alpha\}$的局部有限性,故 U 也与有限多个成员 $U_{\alpha_1},\cdots,U_{\alpha_r}$ 相交,因此(6.5.7)式限制在 U 上成为

$$ds^2 = \sum_{j=1}^{r} h_{\alpha_j} \cdot ds_{\alpha_j}^2 = g_{ij}\,du_i du_j,$$

其中

$$g_{ij} = \sum_{\lambda=1}^{r} \sum_{\mu=1}^{m} h_{\alpha_\lambda} \frac{\partial u_{\alpha_\lambda}^\mu}{\partial u_i} \frac{\partial u_{\alpha_\lambda}^\mu}{\partial u_j}.$$

因为 $0 \leqslant h_\alpha \leqslant 1$, $\sum_\alpha h_\alpha = 1$,故有一个指标 β,使得 $h_\beta(p) > 0$,从而 $ds^2(p) \geqslant h_\beta \cdot ds_\beta^2$,由此得到 ds^2 在 M 上处处正定. 定理得证.

下面是当局部坐标改变$(U,u_j) \to (W,w_j)$时,$G = g_{ij}\,du_i \otimes du_j$ 的分量 g_{ij} 的变换公式:

$$g'_{ij} = g_{i'j'} \frac{\partial u_{i'}}{\partial w_i} \frac{\partial u_{j'}}{\partial w_j}. \tag{6.5.8}$$

例 6.5.1 欧氏空间$\mathbb{R}^n$ 中的弧长微元.

因为$\mathbb{R}^n$ 中的局部坐标就是$(\mathbb{R}^n,(x_1,\cdots,x_n))$,故弧长微元为

$$\mathrm{d}s^2=(\mathrm{d}x_1)^2+\cdots+(\mathrm{d}x_n)^2.$$

6.5.2 连络

矢量丛的截面是流形上的矢量场,我们要对矢量场进行微分,就必须在矢量丛上引进一种新的结构,称为"连络"结构,例如仿射连络、标架丛上的连络、Riemann 连络等. 总之,有了连络概念,能使我们对一个场引进微分结构.

1. 光滑流形的矢量丛 (E,M,π)

把张量丛的构造抽象化,就得到矢量丛的概念.

1) 矢量丛的定义

定义 6.5.4(矢量丛) 给定两个光滑流形 E,M,$\dim E=n$,$\dim M=m$. 设 $\pi:E\to M$ 为光滑映射;又设 V 是 l 维线性空间,$\dim V=l$;并设$\{U,W,\cdots\}$是 M 的一个开覆盖,$\{\varphi_U,\varphi_W,\cdots\}$是一组光滑映射,使得下列条件成立:

(1) $\forall\varphi_U:U\times\mathbb{R}^l\to\pi^{-1}(U)\subset E$ 是可微同胚,且$\forall(p,y)\in U\times\mathbb{R}^l$,有 $\pi\circ\varphi_U(p,y)=p$;

(2) $\forall p\in U$,令 $\varphi_{U,p}(y)=\varphi_U(p,y)$,$y\in\mathbb{R}^l$,则 $\varphi_{U,p}:\mathbb{R}^l\to\pi^{-1}(p)$是一个同胚. 当 $U,W\in\{U,W,\cdots\}$,$U\cap W\neq\varnothing$ 时,映射 $\varphi_{U,W}(p)\equiv\varphi_{W,p}^{-1}\circ\varphi_{U,p}:\mathbb{R}^l\to\mathbb{R}^l$ 是线性空间 $V\leftrightarrow\mathbb{R}^l$ 上的线性自同构,亦即 $\varphi_{U,W}(p)\equiv\varphi_{W,p}^{-1}\circ\varphi_{U,p}\in GL(V)=GL(\mathbb{R}^l)$;

(3) 对于 $U\cap W\neq\varnothing$,映射 $\varphi_{U,W}:U\cap W\to GL(V)$是光滑映射,

则称(E,M,π)**是流形 M 上的(实)n 维矢量丛(vector bundle)**,E **称为丛空间(bundle space)**,M 称为**底空间(underlying space)**,π 称为**丛投影(bundle project)**,V 称为**纤维型(bundle type)**.

进而,$\forall p\in M$,令 $E_p=\pi^{-1}(p)$,称 E_p **为矢量丛 E 在点 $p\in M$ 上的纤维(bundle)**.

$p\in M$ 的张量丛 $T_s^r=\bigcup\limits_{p\in M}T_s^r(p)$ 是矢量丛(E,M,π)的特例,$E\leftrightarrow T_s^r$,图形也与前面张量丛的图形类似.

注 关于上述定义中的条件(2)与条件(3),我们作以下说明:

由(2)得到,纤维型 V 上的元 y_U,y_W 使得充要条件

$$\varphi_U(p,y_U)=\varphi_W(p,y_W)\Leftrightarrow y_U\circ g_{U,W}(p)=y_W$$

成立,其中 $g_{U,W}(p)\in L(\mathbb{R}^l)$视为 $l\times l$ 非退化方阵. 于是,设 U 是M 中包含点 $p\in M$ 的坐标域,则纤维型 V 上的结构可以通过映射 $g_{U,W}(p)\in L(\mathbb{R}^l)$移植到纤维 E_p 上,使得纤维 E_p 成为一个 l 维线性空间,并且这个线性空间结构与点的坐标卡(U,φ_U)的选取无关. 因此,直观上可以把矢量丛(E,M,π)看作诸如 $U\times\mathbb{R}^l$ 的积流形沿着同一点 $p\in M$ 上的纤维 E_p 粘合起来的结果,在粘合时,要求纤维上的线性关系保持不变. 积流形 $U\times\mathbb{R}^l=E$ 是矢量丛最简单的例.

由(3)得到，映射 $\varphi_{U,W}:U\cap W\to GL(V)$ 适合下列相容条件：

① $\forall p\in U$, $\varphi_{U,U}(p)=I:V\to V$；

② 若 $p\in U\cap W\cap Q\neq\varnothing$，则 $\varphi_{U,W}(p)\circ\varphi_{W,Q}(p)\circ\varphi_{Q,U}(p)=I:V\to V$.

于是，把 $\{\varphi_{U,W}\}$ 称为矢量丛 (E,M,π) 的**过渡函数族**，而相容条件①、②是使得 $\{\varphi_{U,W}\}$ 成为过渡函数族的充分条件. 确切地说，有如下定理：

定理 6.5.2(矢量丛的存在性定理)　设 M 为 m 维光滑流形，$\{U_\alpha\}_{\alpha\in J}$ 是 M 的一个开覆盖，V 是 l 维线性空间，若对于每一对指标 $\alpha,\beta\in J$，在 $U_\alpha\cap U_\beta\neq\varnothing$ 时，都确定一个光滑映射 $g_{\alpha,\beta}:U_\alpha\cap U_\beta\to GL(V)$，适合相容条件①、②，则存在 M 上的 l 维矢量丛 (E,M,π)，它以 $\{g_{\alpha,\beta}\}$ 为过渡函数族.

2) 矢量丛的例子

例 6.5.2　矢量丛的对偶丛

设 M 是 m 维光滑流形，$\{(U,\varphi_U)\}$ 是一个给定的局部坐标系，(E,M,τ) 是 M 上的矢量丛；V^* 是 l 维线性空间 V 的对偶空间，E^* 是 M 上的、以 V^* 为纤维型的矢量丛，丛投影记为 π^*，丛的局部积结构为 $\{(U,\psi_U)\}$. 如果对于任一点 $p\in U\cap W\neq\varnothing$，当 $y_U,y_W\in V$ 且 λ_U, $\lambda_W\in V^*$ 分别适合 $\varphi_U(p,y_U)=\varphi_W(p,y_W)$ 与 $\psi_U(p,\lambda_U)=\varphi_W(p,\lambda_W)$ 时，总成立 $\langle y_U,\lambda_U\rangle=\langle y_W,\lambda_W\rangle$，则可定义纤维 $\pi^{-1}(p)$ 与 $(\pi^*)^{-1}(p)$ 的配合

$$\langle\varphi_U(p,y_U),\psi_U(p,\lambda_U)\rangle=\langle y_U,\lambda_U\rangle, \tag{6.5.9}$$

使得两个纤维 $\pi^{-1}(p)$ 与 $(\pi^*)^{-1}(p)$ 互为对偶矢量空间，这时，称两个矢量丛 (E,M,π) 与 (E^*,M,π^*) 互为**对偶丛(dual bundle)**.

若分别取 l 维线性空间 V 与其对偶空间 V^* 的基为对偶基，并把 $y\in V$ 表示为 l 维行向量，$\lambda\in V^*$ 表示为 l 维列向量，则 V 与 V^* 中元的配合可表示为方阵的乘法 $\langle y,\lambda\rangle=y\cdot\lambda$，其中 · 是 $GL(V)$ 相应的方阵的乘法.

由于(6.3.5)式中的 $y_W=y_U\cdot g_{U,W}(p)$，我们有

$$\langle y_U,\lambda_U\rangle=\langle y_W,\lambda_W\rangle\Leftrightarrow y_U\cdot\lambda_U=y_W\cdot\lambda_W,$$

于是，将 $y_W=y_U\cdot g_{U,W}(p)$ 代入上式右边，得 $y_U\cdot\lambda_U=y_U\cdot g_{U,W}(p)\cdot\lambda_W$，所以

$$\lambda_U=g_{U,W}(p)\cdot\lambda_W.$$

类似地可得到对偶丛 E^* 的过渡函数族 $\{h_{U,W}\}$ 中的函数为 $h_{U,W}=([g_{U,W}]^{-1})^{\mathrm{T}}=[g_{W,U}]^{\mathrm{T}}$.

最后，由于切丛 $T(M)$、余切丛 $T^*(M)$ 都是矢量丛的特例，所以，当 (E,M,π) 是 M 的切丛时，过渡函数族 $\{k_{U,W}\}$ 由坐标变换的 Jacobi 方阵组成，(E^*,M,π^*) 则成为 M 的余切丛，余切丛的过渡函数族中的函数恰是 $k_{U,W}$ 的转置逆方阵，所以余切丛恰是切丛的对偶丛.

例 6.5.3　两个矢量丛 E 与 E' 的直和 $E\oplus E'$.

设 E 与 E' 都是流形 M 上的矢量丛，纤维型分别是 V 与 V'，过渡函数族分别为 $\{g_{U,W}\}$ 与 $\{g'_{U,W}\}$. 令

$$h_{U,W}=\begin{bmatrix} g_{U,W} & 0 \\ 0 & g'_{U,W} \end{bmatrix},$$

则$\{h_{U,W}\}$是$V\oplus V'$上的线性自同构，且$\{h_{U,W}\}$适合过渡函数族的相容条件①、②.

流形 M 上的以$\{h_{U,W}\}$为过渡函数族、以 $V\oplus V'$为纤维型的矢量丛，称为 E 与 E'的**直和(direct sum)**，记为 $E\oplus E'$.

例 6.5.4 两个矢量丛 E 与 E'的张量积 $E\otimes E'$.

设 E 与 E'都是流形 M 上的矢量丛，纤维型分别是 V 与 V'，过渡函数族分别为$\{g_{U,W}\}$与$\{g'_{U,W}\}$. 令 $k_{U,W}$是 $g_{U,W}$与 $g'_{U,W}$的张量积，亦即 $k_{U,W}=g_{U,W}\otimes g'_{U,W}$，张量积 $V\otimes V'$在 $k_{U,W}$上的作用定义为

$$(v\otimes v')\cdot k_{U,W}=(v\cdot g_{U,W})\otimes(v'\cdot g'_{U,W}),\quad v\in V, v'\in V'.$$

于是，流形 M 上以$\{k_{U,W}\}=\{g_{U,W}\otimes g'_{U,W}\}$为过渡函数族，以 $V\otimes V'$为纤维型的矢量丛，称为 E 与 E'的张量积，记为 $E\otimes E'$.

3）矢量丛的光滑截面

流形 M 上的(r,s)型张量丛 $T^r_s\equiv T^r_s(M)$是 r 个切丛与 s 个余切丛的张量积. 因此，类似于定义 6.3.4，可定义矢量丛的光滑截面.

定义 6.5.5(矢量丛的光滑截面) 设 M 为 $\dim M=m$ 光滑流形，(E,M,π)为 M 的矢量丛，如果映射 $t:M\to E$ 是光滑映射，使得 $\pi\circ t\equiv I:M\to M$，则称 $t:M\to E$ 是**矢量丛**(E,M,π)**的一个光滑截面**，记矢量丛(E,M,π)的光滑截面的全体为 $\Gamma(E)$.

因为矢量丛的每一条纤维 $E_p=\pi^{-1}(p)$是与纤维型 V 同构的线性空间，所以可以在光滑截面全体所成的集合 $\Gamma(E)$上定义“两截面的加法”、“截面与数的乘法”：

$$(t_1+t_2)(p)=t_1(p)+t_2(p),\quad t_1,t_2\in\Gamma(E);$$

$$(\alpha t)(p)=\alpha(p)t(p),\quad t\in\Gamma(E),$$

其中 α 是流形 M 上的光滑实值函数. 于是，$\Gamma(E)$是一个实线性空间.

若 E 取为切丛 $T(M)$，则光滑截面空间 $\Gamma(T(M))$就是光滑切矢量场 X(定义 6.3.6)所成的线性空间，且 $X\in\Gamma(T(M))\Leftrightarrow$对于 $p\in M$，有 $X\in T_p(M)$.

注意到，矢量丛(E,M,π)的处处不为零的光滑截面不一定存在. 实际上，不为零的光滑截面的存在性反映了光滑流形的一定的拓扑性质.

2. 矢量丛上的连络

设 E 是光滑流形 M 上的一个 q 维实矢量丛，$\Gamma(E)$是矢量丛 E 在流形 M 上的光滑截面的集合. 其实，$\Gamma(E)$是一个实线性空间，也是一个 $C^\infty(M)$-模(模的概念见第 1 章).

定义 6.5.6(连络) 设 E 为矢量丛，E 上的一个映射 D 称为一个**连络(connection)**，若

$$\mathrm{D}:\Gamma(E)\to\Gamma(T^*(M)\otimes E)\tag{6.5.10}$$

满足

(1) 对于任意 $s_1, s_2 \in \Gamma(E)$，有 $\mathrm{D}(s_1 + s_2) = \mathrm{D}s_1 + \mathrm{D}s_2$；

(2) 对于 $s \in \Gamma(E)$ 及任意的 $\alpha \in C^\infty(M)$，有 $\mathrm{D}(\alpha s) = \mathrm{d}\alpha \otimes s + \alpha \mathrm{D}s$.

注 1 若取 $\alpha = -1$，则得 $\mathrm{D}(-s) = -\mathrm{D}s$，所以，D 把零截面映到零截面.

注 2 D 是 $\Gamma(E)$ 到 $\Gamma(T^*(M) \otimes E)$ 的线性算子，$T^*(M)$ 是光滑流形 M 的余切丛.

注 3 D 是作用在 E 的截面上的算子，但它具有局部性：若 $s_1, s_2 \in \Gamma(E)$ 是两个截面，而且限制在 M 的一个开集 U 上相同，则 $\mathrm{D}s_1, \mathrm{D}s_2$ 限制在 U 上也相同.

利用 D 的局部性，可以把 D 定义为作用在局部截面上的算子. 因此，$\mathrm{D}s$ 是在 U 上完全确定的截面.

定理 6.5.3(连络的存在性定理) 在光滑流形 M 的任一个矢量丛 (E, M, π) 上，连络总是存在的. 进而，切丛 $T(M)$ 作为光滑流形 M 的一个特殊的矢量丛，其上的连络一定存在.

证明可参看[3].

3. 绝对微商

定义 6.5.7(绝对微商) 设 X 是光滑流形 M 上的光滑切矢量场，$X \in \Gamma(T(M))$；又 (E, M, π) 是 M 上的矢量丛，$\Gamma(E)$ 是矢量丛 E 的光滑截面空间，D 是 E 上的一个连络. 对于截面 $s \in \Gamma(E)$，令

$$\mathrm{D}_X s = \langle X, \mathrm{D}s \rangle, \tag{6.5.11}$$

其中记号 $\langle \cdot, \cdot \rangle$ 为 $T(M)$ 与 $T^*(M)$ 之间的配合，则 $\mathrm{D}_X s$ 是 E 的截面，即 $\mathrm{D}_X s \in \Gamma(E)$，称 $\mathrm{D}_X s$ 为**截面 s 沿切矢量场 X 的绝对微商(absolute derivative)**.

定理 6.5.4(绝对微商的性质) 设 M 是 m 维光滑流形，由于 $\mathrm{D}_X s = \langle X, \mathrm{D}s \rangle$，故连络 D 可视为二元映射：

$$\forall X \in \Gamma(T(M)), \forall s \in \Gamma(E) \Rightarrow \mathrm{D}: \Gamma(T(M)) \times \Gamma(E) \to \Gamma(E);$$

则 D_X 具有以下性质：当 $X, Y \in \Gamma(T(M)), s, s_1, s_2 \in \Gamma(E), \alpha \in C^\infty(M)$ 时，有

(1) $\mathrm{D}_{X+Y} s = \mathrm{D}_X s + \mathrm{D}_Y s$；

(2) $\mathrm{D}_{\alpha X} s = \alpha \mathrm{D}_X s$；

(3) $\mathrm{D}_X(s_1 + s_2) = \mathrm{D}_X s_1 + \mathrm{D}_X s_2$；

(4) $\mathrm{D}_X(\alpha s) = X\alpha + \alpha \mathrm{D}_X s$.

定理 6.5.5(绝对微商的局部性质) 若 X_1, X_2 是光滑流形 M 上在点 $p \in M$ 取值相同的两个光滑切矢量场，则对于 E 的任一个截面 $s \in \Gamma(E)$，绝对微商 $\mathrm{D}_{X_1} s$ 与 $\mathrm{D}_{X_2} s$ 在 p 点也取相同的值.

由此定理，我们得到：

① 定义 E 的光滑截面空间 $\Gamma(E)$ 关于 M 在点 $p \in M$ 的切矢量场 X 的绝对微商为

$$\forall X \in T_p(M) \Rightarrow \mathrm{D}_X: \Gamma(E) \to E_p.$$

② 对于点 $p\in M$，映射 $\mathrm{D}_X:\Gamma(E)\to E_p$，$\forall X\in T_p(M)$，只要截面 s_1,s_2 在 M 的一条与 X 相切的参数曲线上取值相同，就有 $\mathrm{D}_X s_1=\mathrm{D}_X s_2$.

4. 局部标架场，连络方阵，曲率方阵

定义 6.5.8（局部标架场） 设 $\{(U,u_j)\}$ 是光滑流形 M 的一个局部坐标系，取矢量丛 E 在 U 上的 q 个光滑截面 $s_\alpha\in\Gamma(E)$，$1\leqslant\alpha\leqslant q$，使得它们处处线性无关，称这 q 个光滑截面为 E 在 U 上的一个**局部标架场（local frame field）**，记为 $\{s_\alpha,1\leqslant\alpha\leqslant q\}$.

对于每个点 $p\in U$，可以利用局部标架场 $\{s_\alpha,1\leqslant\alpha\leqslant q\}$，给出张量空间 $T_p^*\otimes E_p$ 的基
$$\{\mathrm{d}u_j\otimes s_\alpha:1\leqslant j\leqslant m,1\leqslant\alpha\leqslant q\}.$$
因为 $\mathrm{D}s_\alpha$ 是矢量丛 $T^*(M)\otimes E$ 在 U 上的局部截面，可令
$$\mathrm{D}s_\alpha=\sum_{i=1}^{m}\sum_{\beta=1}^{q}\Gamma_{\alpha,i}^{\beta}\mathrm{d}u_i\otimes s_\beta,\tag{6.5.12}$$
其中 $\Gamma_{\alpha,i}^{\beta}$ 是 U 上的光滑函数，于是，记 $\omega_\alpha^\beta=\sum_{i=1}^{m}\Gamma_{\alpha,i}^{\beta}\mathrm{d}u_i$，使得(6.5.12)式写为
$$\mathrm{D}s_\alpha=\sum_{\beta=1}^{q}\omega_\alpha^\beta\otimes s_\beta.\tag{6.5.13}$$
将上式写为矩阵形式：令 $S=\begin{bmatrix}s_1\\ \vdots\\ s_q\end{bmatrix}$，$\omega=[\omega_\alpha^\beta]_{1\leqslant\alpha,\beta\leqslant q}=\begin{bmatrix}\omega_1^1&\cdots&\omega_1^q\\ \vdots&\ddots&\vdots\\ \omega_q^1&\cdots&\omega_q^q\end{bmatrix}$ 分别表示局部标架场所成的列矩阵与 ω_α^β 所成的方阵，称 $\omega=[\omega_\alpha^\beta]_{1\leqslant\alpha,\beta\leqslant q}$ 为**连络方阵（connection matrix）**. 于是(6.5.13)式可记为矩阵的张量积 $\mathrm{D}S=\omega\otimes S$. 这里矩阵 $A=[a_{ij}]_{m\times n}$ 与 $B=[b_{kl}]_{r\times s}$ 的张量积定义为
$$A\otimes B=\begin{bmatrix}a_{11}B&\cdots&a_{1n}B\\ \vdots&\ddots&\vdots\\ a_{m1}B&\cdots&a_{mn}B\end{bmatrix}.$$
设另有一个局部标架场 $S'=\begin{bmatrix}s_1'\\ \vdots\\ s_q'\end{bmatrix}$，它与原局部标架场 $S=\begin{bmatrix}s_1\\ \vdots\\ s_q\end{bmatrix}$ 的关系是 $S'=AS$，其中 $\det A\neq0$，$A=\begin{bmatrix}a_1^1&\cdots&a_1^q\\ \vdots&\ddots&\vdots\\ a_q^1&\cdots&a_q^q\end{bmatrix}$，且 a_j^k 是 U 上的光滑函数. 则连络方阵 ω 在局部标架场改变时，$\omega\to\omega'$ 有重要的变换公式：
$$\omega'=\mathrm{d}A\cdot A^{-1}+A\cdot\omega\cdot A^{-1}.\tag{6.5.14}$$
改写(6.5.14)式为 $\omega'\cdot A=\mathrm{d}A+A\cdot\omega$，两边求一次外微分，得到
$$\mathrm{d}\omega'\cdot A-\omega' ?\mathrm{d}A=\mathrm{d}A ?\omega+A\cdot\mathrm{d}\omega,\tag{6.5.15}$$

其中方阵之间的外积?表示方阵在相乘时，元素的积是外积. 因此将 $dA=\omega'\cdot A-A\cdot\omega$ 代入(6.5.15)式，得

$$(d\omega'-\omega'?\omega')\cdot A=A\cdot(d\omega-\omega?\omega). \tag{6.5.16}$$

定理 6.5.6(局部标架场的存在性定理) 设 D 是矢量丛 E 上的一个连络，$p\in M$，则在点 p 的一个坐标域 U 上，存在局部标架场 S，使得对应的连络方阵 ω 在点 p 为零.

定义 6.5.9(曲率方阵) 称 $\Omega=d\omega-\omega?\omega$ 为连络 D 在 U 上的**曲率方阵(curvature matrix)**.

由此，曲率方阵在局部标架场改变时的变换公式(6.5.16)可改写为

$$\Omega'=A\cdot\Omega\cdot A^{-1}. \tag{6.5.17}$$

注 曲率方阵 $\Omega=[\Omega_\alpha^\beta]_{1\leqslant\alpha,\beta\leqslant q}$ 的变换公式(6.5.17)是齐次的，而连络方阵 $\omega=[\omega_\alpha^\beta]_{1\leqslant\alpha,\beta\leqslant q}$ 的变换公式(6.5.16)却不是齐次的.

定义 6.5.10(曲率算子) 对于光滑流形 M 上的两个光滑切矢量场 $X,Y\in\Gamma(T(M))$，若 $\forall s\in\Gamma(E)$，$\forall p\in M$，算子 $R(X,Y):\Gamma(E)\to\Gamma(E)$ 满足

$$(R(X,Y)s)(p)=R(X_p,Y_p)s_p, \tag{6.5.18}$$

则称 $R(X,Y)$ 为**连络 D 的曲率算子(curvature operator of connection D)**.

连络 D 的曲率算子 $R(X,Y)$ 的表示：设 $X,Y\in\Gamma(T(M))$ 为光滑流形 M 上的两个光滑切矢量场. 对于 M 上的连络 D 及其曲率方阵 $\Omega=[\Omega_\alpha^\beta]_{1\leqslant\alpha,\beta\leqslant q}$，任取 $p\in U\subset M$，则 $X,Y\in T_p(M)$，故 $\langle X?Y,\Omega_\alpha^\beta\rangle$ 是纤维 $\pi^{-1}(p)$(线性空间)上的(1,1)型张量. 取 $s\in\pi^{-1}(p)$，并将 s 用矢量丛 E 在 U 上的局部标架场 $S_U=\begin{bmatrix}s_1\\ \vdots\\ s_q\end{bmatrix}$ 表示，得到 $s=\sum\limits_{\alpha=1}^{q}\lambda_\alpha(s_\alpha)\big|_p$，$\lambda_\alpha\in\mathbb{R}$. 于是，算子

$$R(X,Y)s\equiv\sum_{\alpha=1}^{q}\sum_{\beta=1}^{q}\lambda_\alpha\langle X?Y,\Omega_\alpha^\beta\rangle(s_\beta)\big|_p \tag{6.5.19}$$

与局部标架场的选取无关，而由连络 D 与曲率方阵 Ω 决定，且它是 $\pi^{-1}(p)$ 到自身的线性映射. 因此所定义的曲率算子有意义.

定理 6.5.7 曲率算子 $R(X,Y)$ 有如下性质：

(1) $R(X,Y)=-R(Y,X)$；

(2) $R(fX,Y)=f\cdot R(X,Y)$；

(3) $R(X,Y)(fs)=f\cdot(R(X,Y)s)$，

其中 $X,Y\in\Gamma(T(M))$，$f\in C^\infty(M)$，$s\in\Gamma(E)$.

定理 6.5.8 曲率算子 $R(X,Y)$ 是局部算子. 用连络 D 与绝对微商，可表示为

$$R(X,Y)=D_XD_Y-D_YD_X-D_{[X,Y]}.$$

关于“平行”概念，我们可以推广到流形上.

定义 6.5.11(平行截面) 若矢量丛 E 的截面 $s\in\Gamma(E)$ 满足条件 $\mathrm{D}s=0$，则称 s 为**平行截面(parallel section)**.

设 $\gamma\subset M$ 是光滑流形 M 上的一条光滑参数曲线，X 是 γ 的切矢量场. 若矢量丛 E 在 γ 上的截面 s 满足 $\mathrm{D}_X s=0$，则称 s 沿曲线 γ 是平行的.

5. 仿射连络

m 维光滑流形 M 上的切丛 $T(M)$ 是一个由 M 的微分结构本身所确定 m 维矢量丛；切丛 $T(M)$ 上的连络一定存在(见定理 6.5.3)，这种连络称为仿射连络.

定义 6.5.12(仿射连络，仿射连络空间，容许连络) 设 M 为 m 维光滑流形，$T(M)$ 为 M 的切丛，称 $T(M)$ 上的连络为流形 M 上的**仿射连络(affine connection)**.

光滑流形 M 上的仿射连络一定存在. 若光滑流形 M 上给定一个仿射连络，则称 M 为**仿射连络空间.**

设 (M,G) 是 m 维广义 Riemann 流形，D 是 M 上的一个仿射连络，若 $\mathrm{D}G=0$，则称仿射连络 D 是广义 Riemann 流形 (M,G) 的**容许仿射连络(admissible affine connection)**，简称容许连络.

6. 矢量场的绝对微分

取 m 维光滑流形 M 的一个局部坐标系 $\{(U,u_j)\}$，则 $p\in U$ 的切空间 T_p 的自然基 $\left\{\frac{\partial}{\partial u_i},1\leqslant i\leqslant m\right\}$ 构成切丛 $T(M)$ 在 U 上的局部标架场 $\left\{s_i=\frac{\partial}{\partial u_i},1\leqslant i\leqslant m\right\}$，因此可设

$$\mathrm{D}s_i=\sum_{j=1}^{m}\omega_i^j\otimes s_j\equiv\sum_{k=1}^{m}\sum_{j=1}^{m}\Gamma_{i,k}^{j}\mathrm{d}u_k\otimes s_j,\quad 1\leqslant i\leqslant m,\tag{6.5.20}$$

其中 $\Gamma_{i,k}^{j}$ 是 U 上的光滑函数，称 $\Gamma_{i,k}^{j}$ 为连络 D 关于局部坐标 u_i 的系数，简称**连络系数.**

定理 6.5.9 当局部坐标系变换时，$\{(U,u_j)\}\to\{(W,w_j)\}$，局部标架场变化为

$$S=\left\{s_i=\frac{\partial}{\partial u_i},1\leqslant i\leqslant m\right\}\to S'=\left\{s_i'=\frac{\partial}{\partial w_i},1\leqslant i\leqslant m\right\},$$

则 $S'=J_{WU}\cdot S$，其中 J_{WU} 是局部坐标变换的 Jacobi 行列式，有

$$J_{WU}=\begin{bmatrix}\frac{\partial u_1}{\partial w_1} & \cdots & \frac{\partial u_m}{\partial w_1}\\ \vdots & \ddots & \vdots\\ \frac{\partial u_1}{\partial w_m} & \cdots & \frac{\partial u_m}{\partial w_m}\end{bmatrix}.$$

从而，也得到连络系数 $\Gamma_{i,k}^{j}$ 的坐标变换公式

$$\Gamma'^{j}_{i,k}=\sum_{i',j',k'=1}^{m}\Gamma_{i',k'}^{j'}\frac{\partial w_j}{\partial u_{j'}}\frac{\partial u_{i'}}{\partial w_i}\frac{\partial u_{k'}}{\partial w_k}+\sum_{i'=1}^{m}\frac{\partial^2 u_{i'}}{\partial w_i\partial w_k}\cdot\frac{\partial w_j}{\partial u_{i'}}.\tag{6.5.21}$$

注 (6.5.21)式表明,连络系数 $\Gamma_{i,k}^{j}$ 不满足 M 上的张量分量的变化规律.

定义 6.5.13(切矢量场的绝对微分) 设 M 为 m 维光滑流形,X 是 M 上的光滑切矢量场,$X=\sum_{i=1}^{m} x_i \frac{\partial}{\partial u_i}$;D 是 M 上给定的仿射连络. 称 $\mathrm{D}X=\sum_{i,j=1}^{m}(\mathrm{d}x_i+x_j\omega_j^i)\otimes\frac{\partial}{\partial u_j}$ 为 X 的绝对微分,其中 ω_j^i 是连络方阵 ω 的元. 绝对微分是矢量丛 $T^*(M)\otimes T(M)$ 的截面,**它是流形 M 上的(1,1)型张量场**. 数量场的绝对微分规定为它的普通微分.

7. 曲率张量、挠率张量

对于 M 上的连络 D 及其曲率方阵 $\Omega=[\Omega_i^j]_{1\leqslant i,j\leqslant q}$,连络方阵 $\omega=[\omega_i^j]_{1\leqslant i,j\leqslant q}$,则

$$\Omega_i^j = \sum_{k=1}^{m}\Gamma_{i,k}^{j}\mathrm{d}u_k,\quad 1\leqslant i,j\leqslant m, \tag{6.5.22}$$

$\Gamma_{i,k}^{j}$ 是 U 上的光滑函数,亦即连络 D 关于局部坐标 u_j 的系数. 故曲率方阵 $\Omega=\mathrm{d}\omega-\omega?\omega$ 可表示为

$$\begin{aligned}
\mathrm{d}\omega_i^j-\omega_i^h?\omega_h^j &= \frac{\partial\Gamma_{i,k}^{j}}{\partial u_l}\mathrm{d}u_l\,?\mathrm{d}u_k-\Gamma_{i,l}^{h}\Gamma_{h,k}^{j}\mathrm{d}u_l\,?\mathrm{d}u_k\\
&=\frac{1}{2}\left(\frac{\partial\Gamma_{i,l}^{j}}{\partial u_k}-\frac{\partial\Gamma_{i,k}^{j}}{\partial u_l}+\Gamma_{i,l}^{h}\Gamma_{h,k}^{j}-\Gamma_{i,k}^{h}\Gamma_{h,l}^{j}\right)\mathrm{d}u_k?\mathrm{d}u_l\\
&\equiv\frac{1}{2}R_{i,k,l}^{j}\mathrm{d}u_k?\mathrm{d}u_l,
\end{aligned}$$

其中 $R_{i,k,l}^{j}=\frac{\partial\Gamma_{i,l}^{j}}{\partial u_k}-\frac{\partial\Gamma_{i,k}^{j}}{\partial u_l}+\Gamma_{i,l}^{h}\Gamma_{h,k}^{j}-\Gamma_{i,k}^{h}\Gamma_{h,l}^{j}$.

定义 6.5.14(曲率张量) 称 $R=R_{i,k,l}^{j}\frac{\partial}{\partial u_j}\otimes\mathrm{d}u_i\otimes\mathrm{d}u_k\otimes\mathrm{d}u_l$ 为 m **维光滑流形 M 的仿射连络 D 的曲率张量.**

曲率张量不依赖于局部坐标系的选取,当$\{(U,u_j)\}\to\{(W,w_j)\}$时,则

$$R_{i,k,l}^{j} = R_{i',k',l'}^{j'}\frac{\partial w_j}{\partial u_{j'}}\frac{\partial u_{i'}}{\partial w_i}\frac{\partial u_{k'}}{\partial w_k}\frac{\partial u_{l'}}{\partial w_l}, \tag{6.5.23}$$

它适合(1,3)型张量的分量变换规律,因此**曲率张量是流形 M 上的一个(1,3)型张量场.**

定理 6.5.10 设 X,Y,Z 为三个光滑切矢量场,其局部表示分别为 $X=X_i\frac{\partial}{\partial u_i}$,$Y=Y_i\frac{\partial}{\partial u_i}$,$Z=Z_i\frac{\partial}{\partial u_i}$,则曲率算子 $R(X,Y)$ 用曲率张量表示为 $R(X,Y)Z=Z_i\langle X\,?Y,\Omega_i^j\rangle\frac{\partial}{\partial u_i}=R_{i,k,l}^{j}Z_iX_kY_l\frac{\partial}{\partial u_i}$,因此,曲率张量 R 的"系数"$R_{i,k,l}^{j}$ 有表示

$$R_{i,k,l}^{j} = \left\langle R\left(\frac{\partial}{\partial u_k},\frac{\partial}{\partial u_l}\right)\frac{\partial}{\partial u_i},\mathrm{d}u_j\right\rangle. \tag{6.5.24}$$

定义 6.5.15(挠率张量) 对于连络系数 $\Gamma^j_{i,k}$,令 $T^j_{i,k}=\Gamma^j_{k,i}-\Gamma^j_{i,k}$,则

$$T=T^j_{i,k}\frac{\partial}{\partial u_j}\otimes \mathrm{d}u_i\otimes \mathrm{d}u_k$$

称为 m **维光滑流形** M **的仿射连络** D **的挠率张量(torsion tensor)**.

挠率张量不依赖于局部坐标的选取,当$\{(U,u_j)\}\to\{(W,w_j)\}$,则 $T^j_{i,k}=T^{j'}_{i',k'}\frac{\partial w_j}{\partial u_{j'}}\frac{\partial u_{i'}}{\partial w_i}\frac{\partial u_{k'}}{\partial w_k}$,它适合(1,2)型张量的分量变换规律,因此**挠率张量是流形** M **上的一个**(1,2)**型张量场**.

定义 6.5.16(无挠连络) 若仿射连络 D 的挠率张量为零,则称该连络是**无挠仿射连络(torsionless affine connection)**,简称无挠连络.

8. 几个重要的存在性定理

定理 6.5.1 是 m 维光滑流形 M 上 Riemann 度量的存在性定理;定理 6.5.2 是光滑流形 M 上的矢量丛 E 的存在性定理;定理 6.5.3 是在任一个矢量丛 E 上连络的存在性定理;定理 6.5.6 给出了使得连络方阵为零的局部标架场的存在性定理. 还有两个重要的存在性定理.

定理 6.5.11(无挠仿射连络的存在性定理) m 维光滑流形 M 上必定存在无挠仿射连络. 又设 D 是 M 的无挠仿射连络,则在 M 的任一点 p,存在局部坐标系$\{(U,u_j)\}$,使得对应的连络系数 $\Gamma^j_{i,k}$ 在该点为零.

定理 6.5.12(Riemann 几何的基本定理) 设 M 为 m 维广义 Riemann 流形,则 M 上存在惟一的无挠容许仿射连络.

上述基本定理中存在的无挠容许仿射连络,称为 M 的 Riemann 连络,或 Levi-Civita 连络.

9. 测地线、测地法坐标

定义 6.5.17(流形上的测地线) 设 M 为 m 维光滑流形,$\gamma: u_i=u_i(t),t\in[t_0,T]$,是 M 上的光滑参数曲线,$X(t)$是定义在 γ 上的切矢量场,表示为 $X(t)=\sum_{i=1}^m x_i(t)\left(\frac{\partial}{\partial u_i}\right)_{\gamma(t)}$. 称切矢量场 $X(t)$沿曲线 γ 是平行的,若它沿曲线 γ 的绝对微分为零,亦即 $\mathrm{D}X=0$. 若曲线 γ 的切矢量场沿 γ 自身是平行的,则称 γ 为**测地线(geodesic)**.

对于测地线 γ,它的切矢量场 $X(t)=\sum_{i=1}^m\frac{\mathrm{d}u_i(t)}{\mathrm{d}t}\left(\frac{\partial}{\partial u_i}\right)_{\gamma(t)}$ 沿曲线 γ 是平行的,故测地线 γ 应满足方程组

$$\frac{\mathrm{d}^2u_i}{\mathrm{d}t^2}+\sum_{j,k=1}^m\Gamma^i_{j,k}\frac{\mathrm{d}u_j}{\mathrm{d}t}\frac{\mathrm{d}u_k}{\mathrm{d}t}=0,\quad i=1,2,\cdots,m, \tag{6.5.25}$$

这是二阶常微分方程组，所以通过流形 M 上的任一点，恰有一条测地线，在该点与任意给定的一致切矢量相切.

定义 6.5.18（Riemann 流形上的测地线） 设 M 为 m 维 Riemann 流形，若光滑参数曲线 $\gamma: u_j = u_j(t), t \in [t_0, T]$，是 M 上关于 Levi-Civita 连络的测地线，则称 γ 是 Riemann 流形 M 的测地线.

在欧氏空间 $\mathbb{R}^3$ 中，球面上的测地线（短程线）就是球面上的大圆.

定理 6.5.13 设 M 为 m 维 Riemann 流形，则 M 上的测地线 γ 的参数 t 必是弧长参数 s 的一次函数：$t = \lambda s + \mu$，其中 $\lambda \neq 0, \lambda, \mu$ 为常数.

事实上，设在 M 的局部坐标系 $\{(U, u_j)\}$ 下，Riemann 连络 D 的系数是 $\Gamma_{i,k}^{j}$，则曲线 $\gamma: u_i = u_i(t), 1 \leqslant i \leqslant m$，为测地线的条件是它满足方程组(6.5.25). 按求和约定，可将(6.5.25)式写为

$$\frac{\mathrm{d}^2 u_i}{\mathrm{d}t^2} + \Gamma_{j,k}^{i} \frac{\mathrm{d}u_j}{\mathrm{d}t} \frac{\mathrm{d}u_k}{\mathrm{d}t} = 0, \tag{6.5.26}$$

这里 $\frac{\mathrm{d}u_i}{\mathrm{d}t} = X_i$ 是 γ 的切矢量场. 由定义，测地线的切矢量场沿曲线 γ 本身关于 Riemann 连络是平行的，而 Riemann 连络保持度量性质在平行移动下不变，故测地线的切矢量场的长度是一个常数，亦即 $\frac{\mathrm{d}s}{\mathrm{d}t}$ 为常数，从而得到 $t = \lambda s + \mu$.

于是，可以建立流形 M 上一点 $p \in M$ 附近的一个特殊的坐标系，使得从 p 点出发的任一条测地线的坐标的常微分方程是弧长参数的线性函数，我们称这种坐标系为**测地法坐标系**.

测地法坐标系的建立方法如下：在 $p \in M$ 的局部坐标系 $\{(U, u_j)\}$ 之下，给出任一点 $x \in U$ 的初始值，解方程组

$$\begin{cases} \dfrac{\mathrm{d}^2 u_i}{\mathrm{d}t^2} + \Gamma_{j,k}^{i} \dfrac{\mathrm{d}u_j}{\mathrm{d}t} \dfrac{\mathrm{d}u_k}{\mathrm{d}t} = 0, \\ u_i(0) = x_i, \\ \dfrac{\mathrm{d}u_i}{\mathrm{d}t}(0) = a_i, \end{cases} \tag{6.5.27}$$

得到惟一的一组解

$$u_i = f_i(t, x_k, a_k), \quad |t| < \delta, \tag{6.5.28}$$

其中函数光滑地依赖于其变量 t, x_k, a_k.

经过一定的演算，可以证明：取初值 $a_k = a_k^0$ 为点 $x \in W \subset U$ 的局部坐标，使得 $\{(W, a_k)\}$ 为点 x 的局部坐标系，称为点 x 的**测地法坐标系**，并称 a_k 为 x 的**测地法坐标**，或简称**法坐标**.

由于流形的切空间是线性空间，它上面的坐标系只相差一个非退化的线性变换，因此，流形的法坐标系也只相差一个非退化的线性变换而完全确定. 于是，固定一组 $a_k = a_k^0$，当

t 变化时，ta_k^0 是 $T_x(M)$中从原点出发的一条直线，而流形 M 上则描出一条从 x 出发的、与切矢量(a_k^0)相切的测地线，所以，在测地法坐标系$\{(W,a_k)\}$之下，这条测地线的方程为 $a_k = ta_k^0$，其中 a_k^0 为常数.

定理 6.5.14 设 M 为 m 维无挠仿射连络空间，则对于在每点 $x\in M$ 的测地法坐标系$\{(W,a_j)\}$，其连络系数 $\Gamma_{j,k}^i$在点 x 为零，亦即 $\Gamma_{j,k}^i(0)=0, 1\leqslant i,j,k\leqslant m$.

定理 6.5.15(测地法坐标系的存在性定理) m 维仿射连络空间 M 的每点 $x\in M$，都有一个邻域 W，使得 W 中的每一点都存在包含 W 在内的测地法坐标系$\{(W,a_k)\}$.

测地法坐标系的建立给 Riemann 流形的研究带来很多方便，给出许多有用的性质；读者可参考有关文献.

6.5.3 Lie 群与活动标架法

1. Lie 群

欧氏空间$\mathbb{R}^n$ 的运算结构、拓扑结构、微分结构等，为我们提供了空间结构的丰富模型，而且，当几种结构同时存在时，空间将有更丰富的内涵，本节简单介绍运算结构与微分结构共存时的一种集合——Lie 群.

定义 6.5.19(Lie 群) 设 G 是一个非空集合，若

(1) G 是一个群$(G,\cdot)$；

(2) G 是一个 r 维光滑流形；

(3) 映射 $\tau:G\to G, \tau(g)=g^{-1}$与映射 $\varphi:G\times G\to G, \varphi(g_1,g_2)=g_1g_2$ 都是光滑映射，

则称 G 是一个 r **维 Lie 群(Lie group)**.

例 6.5.5 $\mathbb{R}^n$ 关于矢量的加法成为一个 n 维 Lie 群.

例 6.5.6 环群$\mathbb{T}^n$ 是一个 n 维 Lie 群.

在$\mathbb{R}^n$ 中取基$\{e_i:1\leqslant i\leqslant n\}$，由它们生成的“格”$L=\left\{x=\sum\limits_{i=1}^{n} a_ie_i : a_i\in\mathbb{Z}\right\}=\mathbb{Z}^n$是$\mathbb{R}^n$ 的一个加子群，而“环群”则是商群$\mathbb{T}^n=\mathbb{R}^n/L$. 在拓扑结构上，$\mathbb{T}^n$ 是 n 维环面，因此它是一个紧致 Lie 群.

例 6.5.7 设 G_1,G_2 为两个 Lie 群，在积流形 $G_1\times G_2$ 上定义运算 $\cdot$：

$$(a_1,a_2),(b_1,b_2)\in G_1\times G_2 \Rightarrow (a_1,a_2)\cdot(b_1,b_2)=(a_1b_1,a_2b_2)\in G_1\times G_2,$$

则$(G_1\times G_2,\cdot)$成为一个 Lie 群，称为两个 Lie 群的直积.

例 6.5.8 线性群 $GL(n,\mathbb{R}), L(n,\mathbb{C})$. (见第 1 章)

$GL(n,\mathbb{R})$是 $n\times n$ 阶非退化实方阵所成的集合，群运算是矩阵乘法，拓扑结构、微分结构取$\mathbb{R}^{n^2}$ 的拓扑结构与微分结构，则它构成一个非交换的 Lie 群. $GL(n,\mathbb{C})$类似讨论.

例 6.5.9 $SL(n,\mathbb{R})=\{A\in GL(n,\mathbb{R}):\det A=1\}$，$O(n,\mathbb{R})=\{A\in GL(n,\mathbb{R}):AA^{\mathrm{T}}=I\}$

这里 $SL(n,\mathbb{R})$，$O(n,\mathbb{R})$都是$GL(n,\mathbb{R})$的正则子流形，因而都是 Lie 子群，前者称为幺模群，后者称为(实)正交群.

定义 6.5.20(群的右平移、左平移) 设$(G,\cdot)$是一个群，映射 $R_g:G\to G$ 由 $R_g(x)=x\cdot g$ 给出，则称其为 G 上的**右平移**；映射 $L_g:G\to G$ 由 $L_g(x)=g\cdot x$ 给出，称其为 G 上的**左平移**. 若切矢量场 X 经过右(左)平移后不变，则称其为**右(左)不变切矢量场**.

定理 6.5.16 设 G 为 r 维 Lie 群，若 X,Y 是 G 上的右不变切矢量场，则 Poisson 括号 $[X,Y]=XY-YX$ 也是 G 上的右不变切矢量场；对左不变切矢量场有同样结论.

定义 6.5.21(Lie 代数) 当一个 r 维实线性空间具有满足分配律、反交换律、Jacobi 恒等式的乘法运算时，就称它为一个 **r 维 Lie 代数**.

定义 6.5.22(Lie 群的 Lie 代数结构) 设 G 是一个 Lie 群，记 G 上的右不变光滑切矢量场的全体为 $\mathfrak{A}$ 若右不变光滑切矢量场 $X,Y\in\mathfrak{A}$的 Poisson 括号$[X,Y]$满足

(1) $[c_1X_1+c_2X_2,Y]=c_1[X_1,Y]+c_2[X_2,Y]$， $c_1,c_2\in\mathbb{R}$； (分配律)

(2) $[X,Y]=-[Y,X]$； (反交换律)

(3) $[X,[Y,Z]]+[Y,[Z,X]]+[Z,[X,Y]]=0$， (Jacobi 恒等式)

则称 $\mathfrak{A}$为 **Lie 群 G 的 Lie 代数**.

G 的右不变光滑切矢量场所成的空间 $\mathfrak{A}$具有反交换律、Jacobi 恒等式的“乘法”运算，因此 $\mathfrak{A}$成为一个 Lie 代数.

2. Lie 群的结构常数

设 G 是一个 r 维 Lie 群，e 是其单位元. 右平移 $R_{a^{-1}}:G\to G$ 由 $R_{a^{-1}}(x)=x\cdot a^{-1}$给定，则其满足

(1) $R_{a^{-1}}(e)=e\cdot a^{-1}=a^{-1}$；

(2) $R_{a^{-1}}(a)=a\cdot a^{-1}=e$；

(3) $R_{a^{-1}}:G\to G$ 是 G 到自身的可微同胚.

于是，$R_{a^{-1}}$的切映射$(R_{a^{-1}})_*:T_a(G)\to T_e(G)$是流形 G 在点 $a\in G$ 的切空间 $T_a(G)$到单位元 $e\in G$ 的切空间 $T_e(G)$上的线性同构. 因此，对于 $X\in T_a(G)$，令

$$\omega(X)=(R_{a^{-1}})_*(X), \tag{6.5.29}$$

则 ω 是定义在 $T_a(G)$中、取值在 $T_e(G)$中的一次微分式，称为 Lie 群 G 的**基本微分式**，或 **Maurer-Cartan 形式**.

在 $T_e(G)$中取定基，记为$\{\delta_i,1\leqslant i\leqslant r\}$，则

$$\omega=\sum_{i=1}^{r}\omega^i\delta_i, \tag{6.5.30}$$

其中 ω^i 是 Lie 群 G 上的 r 个处处线性无关的一次微分式.

对于(6.5.30)式的分量 ω^i 求外微分,并令

$$\begin{cases} \mathrm{d}\omega^i = -\dfrac{1}{2}\sum_{j,k=1}^{r} c_{jk}^i \omega^j ? \omega^k, \\ c_{jk}^i + c_{kj}^i = 0. \end{cases} \tag{6.5.31}$$

不难证明,ω^i,$\mathrm{d}\omega^i$ 都是群上的右不变量,故系数 c_{jk}^i 是常数.

定义 6.5.23(Lie 群的结构常数) 称(6.5.31)式中的系数 c_{jk}^i 为 **Lie 群 G 的结构常数.**

定理 6.5.17(Lie 群 G 的结构常数定理) 设 G 为 r 维 Lie 群,则其结构常数满足

$$\sum_{j=1}^{r} (c_{jk}^i c_{hl}^j + c_{jh}^i c_{lk}^j + c_{jl}^i c_{kh}^j) = 0. \tag{6.5.32}$$

通常,在 Lie 群上讨论其左、右不变矢量场、研究其上的各阶微分形式、计算其"结构常数"、Lie 群的 Lie 代数结构、Lie 变换群以及活动标架法等.

3. 活动标架法概述

将 n 维欧氏空间$\mathbb{R}^n$ 的刚体运动群记为$\mathbb{E}(n)$,在$\mathbb{R}^n$ 中取定一个正交标架

$$(\mathbb{R}^n, O; \delta_1, \cdots, \delta_n),$$

将每个刚体运动 $F \in \mathbb{E}(n)$在 $x \in \mathbb{R}^n$ 上的(右)作用记为

$$x \cdot F \equiv F(x) = x \cdot A + a,$$

其中 $x = (x_1, \cdots, x_n) = \sum_{\alpha=1}^{n} x_\alpha \delta_\alpha \in \mathbb{R}^n$,$a = (a_1, \cdots, a_n) = \sum_{\alpha=1}^{n} a_\alpha \delta_\alpha \in \mathbb{R}^n$,方阵 A 是满足 $AA^{\mathrm{T}} = I$,$\det A = 1$ 的正交方阵,$A = \begin{bmatrix} \beta_{11} & \cdots & \beta_{1n} \\ \vdots & \ddots & \vdots \\ \beta_{n1} & \cdots & \beta_{nn} \end{bmatrix}$. 于是,任一刚体运动 F 可用一个"对"(A, a) 与之对应:

$$F \in \mathbb{E}(n) \leftrightarrow (A, a) \in (O(n, \mathbb{R}), \mathbb{R}^n).$$

$\mathbb{E}(n)$中的运算 · 定义为:对于 $F_1, F_2 \in \mathbb{E}(n) \leftrightarrow (A_1, a_1), (A_2, a_2)$,则

$$F_1 \cdot F_2 = (A_1 \cdot A_2,\ a_1 A_2 + a_2),$$

$F \in \mathbb{E}(n) \leftrightarrow (A, a)$的逆定义为

$$(F)^{-1} = (A^{-1}, -aA^{-1}),$$

于是,$\mathbb{E}(n)$是一个$\dfrac{1}{2}n(n+1)$维 Lie 群.

将$\mathbb{R}^n$ 上的正交标架的全体所成的集合记为

$$\mathfrak{F}(\mathbb{R}^n) = \{(\mathbb{R}^n, O; \delta_1, \cdots, \delta_n)\},$$

将 $\mathfrak{F}(\mathbb{R}^n)$中与固定的标架定向一致的正交标架的全体所成的集合记为 $\mathfrak{F}_+(\mathbb{R}^n)$,则 $\mathfrak{F}(\mathbb{R}^n)$,$\mathfrak{F}_+$

$(\mathbb{R}^n)$都是$\mathbb{R}^n$上的主丛，分别以$O(n,\mathbb{R})$，$SO(n,\mathbb{R})$为结构群，于是，$\mathfrak{F}_+(\mathbb{R}^n)\leftrightarrow\mathbb{E}(n)$.

这样，我们建立了$\mathbb{E}(n)$上的活动标架$(\mathbb{E}(n),p;e_1,\cdots,e_n)\leftrightarrow(\mathbb{R}^n,O;\delta_1,\cdots,\delta_n)$，使得$F\in\mathbb{E}(n)\leftrightarrow(A,a)$在活动标架$(\mathbb{E}(n),p;e_1,\cdots,e_n)$下的表示为

$$\begin{cases}\mathrm{d}p=\sum\limits_{\alpha=1}^{n}\omega^{\alpha}e_{\alpha},\\ \mathrm{d}e_{\alpha}=\sum\limits_{\gamma=1}^{n}\omega_{\gamma}^{\alpha}e_{\gamma},\end{cases}\tag{6.5.33}$$

并称ω^{α}，ω_{γ}^{α}，$1\leqslant\alpha,\gamma\leqslant n$为**活动标架的相对分量.**

活动标架的概念起源于力学，在研究刚体运动时，在运动物体上固定一个正交标架，当物体作刚体运动时，正交标架随着运动，得到一族依赖于时间的正交标架，这族标架完全刻划了物体的刚体运动. 后来，法国数学家Cotton、Darboux把单参数活动标架族推广到多参数情形. 于是，活动标架法与外微分相结合，已经成为几何学与物理学的有力工具之一，发挥着巨大作用了.

6.5.4 数学科学与现代物理学

数学大师陈省身先生以微分几何与理论物理为例，精辟地分析了数学科学与现代物理学的关系. 他指出：微分几何与理论物理都是用微积分作为工具，前者是研究几何现象，后者研究物理现象. 后者当然比较广泛些，但任何物理现象都在空间发生，所以前者又是后者的基础. 两者都用推理的方法，理论物理还需有实验的支撑，而几何却不受这个限制，因此，几何，作为数学科学的主要分支之一，选择研究的问题时比较自由，数学科学对研究课题选择的自由度，把数学研究推到新的高度，新的深度与新的领域. 当然，数学的严格性与逻辑性却是必须遵守的.

微分几何与理论物理的关系密切. 例如，广义相对论所需要的Riemann几何理论，规范场理论所需要的矢量丛与连络等概念，都在物理应用前被数学家所发展，后被物理学家成功地应用，给人们以神秘的“殊途同归”之感.

微分几何与理论物理密切联系的第一个重要例子是动力学与活动标架.

在动力学中，描述一个固体运动，是把一个标架安装在固体上，从而用标架的运动来反映固体的运动. 例如，在三维空间$\mathbb{R}^3$中，活动标架是指一点$x\in\mathbb{R}^3$及经过点x的互相垂直的单位矢量e_1,e_2,e_3，如果x也代表该点的坐标矢量，则有

$$\begin{cases}\dfrac{\mathrm{d}x}{\mathrm{d}t}=\sum\limits_{\alpha=1}^{3}p_{\alpha}(t)e_{\alpha},\\ \dfrac{\mathrm{d}e_{\alpha}}{\mathrm{d}t}=\sum\limits_{\gamma=1}^{3}q_{\gamma\alpha}(t)e_{\gamma},\end{cases}$$

这正是(6.5.33)式所表示的，其中t是时间，并且活动标架的分量$q_{\gamma\alpha}$满足$q_{\gamma\alpha}(t)+q_{\alpha\gamma}(t)=0$，于是，函数$p_{\alpha}(t)$，$q_{\gamma\alpha}(t)$完全描述了活动标架与固体的运动.

此外,动力学与空间的曲线论、曲面论有密切联系,且用到了双参数的活动标架方法.我们可以对多参数活动标架法理解为方程组(6.5.33)是偏微分方程组,并将其用外微分的形式表示出来,由方程组

$$\begin{cases} dx = \sum_i \omega_i e_i, \\ de_i = \sum_j \omega_{ji} e_i \end{cases} \quad 与 \quad \omega_{ij} + \omega_{ji} = 0,$$

其中 ω_i 与 ω_{ij} 为参数空间中的一次微分式,求外微分后得

$$\begin{cases} d\omega_i = \sum_j \omega_j ? \omega_{ji}, \\ d\omega_{ij} = \sum_k \omega_{ik} ? \omega_{jk}, \end{cases}$$

这就是著名的 Lie 群 G 的 Maurer-Cartan 方程式. 这个方程式与 Lie 群 G 的 Lie 代数乘法方程式是对偶的,因此从动力学到活动标架,再到 Lie 群的基本方程,是一个非常自然的过程.

以上演变又得到继续推进:爱因斯坦的**广义相对论**发表后,Cartan 继而发表了论文,阐述并发展了**广义仿射空间理论**,并指出这种理论在相对论中的应用,给出了 Maurer-Cartan 方程的推广:

$$\begin{cases} d\omega_i = \sum_j \omega_j ? \omega_{ji}, \\ d\omega_{ij} = \sum_k \omega_{ik} ? \omega_{jk} + \Omega_{ij}, \end{cases}$$

其中 Ω_{ij} 是二次微分式,称为**曲率式**,是三维 Riemann 几何的基本方程.

第二个重要例子是曲面论与孤立子及 σ **模型**.

在三维欧氏空间 $\mathbb{R}^3$ 中,取一个曲面 S,令 x 为 S 的坐标矢量,ξ 表示其单位法矢量,则 S 的不变量是两个二次微分式(参看有关参考书中的"曲面论")

$$\mathrm{I} = (dx, dx) > 0, \quad \mathrm{II} = -(dx, d\xi),$$

分别称为曲面 S 的**第一与第二基本式**. 第一基本式

$$\mathrm{I} = (dx, dx)$$

是正定的,第二基本式的两个特征值

$$\kappa_i, \quad i = 1, 2$$

称为曲面 S 的**主曲率**,对称函数

$$H = \frac{1}{2}(\kappa_1 + \kappa_2) \quad 与 \quad K = \kappa_1 \kappa_2$$

分别称为 S 的中**曲率**与**全曲率**(或 Gauss 曲率).

这些曲率都有其几何意义,例如,使得 $H = \frac{1}{2}(\kappa_1 + \kappa_2) = 0$ 的曲面是最小曲面;又如 $K = -1$ 的曲面,其上的渐近曲线是不重合的实曲线. 若设它们之间的夹角为 φ,则可在 S

上选择参数 u,t，使得 $\varphi_{ut}=\sin\varphi$，这就是著名的 **Sine-Gordon(SG)方程.**

当我们把 $\varphi(u,t)$ 理解为直线 u 上的波动，t 为时间时，则 SG 方程有孤立子解.

还有，常数曲率的曲面 $H=\frac{1}{2}(\kappa_1+\kappa_2)=c$，包括最小曲面，在理论物理上有许多应用，例如可以定义一个调和映射作为能量.

调和映射在物理上还有重要作用. 1980 年物理学家们确定了从二维球面 S^2 到 n 维复投影空间 $P_n(\mathbb{C})$ 的调和映射的全体，称为 σ 模型. 而与最小曲面有关的问题更是既有趣又有用，至今还有许多开问题.

第三个重要例子是规范场论.

规范场论的数学基础是矢量丛的概念. 规范场就是矢量丛的连络.

矢量丛的概念，在近代数学中有决定性的重要性，其要点是把乘积空间 $X\times Y$ 与空间 E 联系起来，考虑映射 $\pi:E\to X$，使得点 $\forall x\in X$ 存在邻域 $U\subset X$，满足条件：$\pi^{-1}(U)$ 与 $U\times Y$ 是拓扑相等的. 显然，$X\times Y$ 与 E 的等同只是局部的. 重要的问题是：这种局部性是否必然是整体性呢？它的解答引导出“示性类”的概念，推动了微分几何的发展.

设 E 为矢量丛，$\pi:E\to X$ 为丛投影. 若映射 $F:X\to E$ 满足条件

$$\pi\circ F(x)=x,\quad x\in X,$$

则称 F 为一个**截面**. 截面的微分需要连络的概念，而量度微分的非交换性，则需要用到曲率.

一切物理的理论最终都要“**量子化**”(**quantization**)，而在数学上，就需要研究无穷维空间及离散(discrete)现象.

还需指出，广义相对论与 Riemann 几何之间的关系密不可分，没有相对论，Riemann 几何就不会受到数学家的重视. 反之，没有 Riemann 几何，相对论就没有数学理论的支撑. 众所周知，Riemann 几何学是德国数学家 Riemann(1826—1866)提出的，是建立在流形上的几何学. 1912 年，爱因斯坦找到了建立广义相对论的数学工具——Riemann 几何学，并于 1915 年提出了描述时空规律的场方程.

习题 6

1. 对于例 3.1.2，找出 $U_2\cap U_3\neq\varnothing$ 所对应的 $\varphi_{U_2},\varphi_{U_3}$，并研究其光滑性.

2. 设 $U\subset\mathbb{R}^3$ 为 $\mathbb{R}^3$ 中的开集，映射 $F:U\to\mathbb{R}$ 为连续可微函数 $F\in C^1(\mathbb{R}^3)$. 记 $(x_1,x_2,x_3)\in\mathbb{R}^3$，若 $F(x_1,x_2,x_3)=ax_1+bx_2+cx_3$，且 a,b,c 是不全为零的实数. $F(x_1,x_2,x_3)=0$ 是 $\mathbb{R}^3$ 中通过原点的平面，它可表示为 $F^{-1}(0)\subset\mathbb{R}^3$. 试证 $F^{-1}(0)$ 是 $\mathbb{R}^3$ 中的二维线性子空间.

3. 证明 m 维光滑流形上的实值函数 $f:M\to\mathbb{R}$ 在点 $p\in M$ 的可微性与容许坐标卡的选取无关.

4. 设 $e_1=(1,0,0),e_2=(1,1,0),e_3=(1,1,1)$ 为 $\mathbb{R}^3$ 的一组基，试求其对偶基 e_1^*,e_2^*,e_3^*.

5. 证明 m 维光滑流形 M 到 n 维光滑流形 N 上的映射 $f:M\to N$ 在 M 上的点 $p\in M$ 的可微性与容许坐标卡的选取无关.

6. 证明定理 6.1.6.

7. 证明关系式(6.1.13)与关系式(6.1.14).

8. 设 M,N 分别为 m,n 维流形，$F:M\to N$ 为光滑映射. 试给出微分映射 $F^*:T_q^*(N)\to T_p^*(M)$ 与切映射 $F_*:T_p(M)\to T_q(N)$ 在自然基之下的变换矩阵，并证明.

9. 试证定理 6.2.2.

10. 理解 $\sigma(x(v_1^*,\cdots,v_r^*))=x(v_{\sigma(1)}^*,\cdots,v_{\sigma(r)}^*)$，$v_j^*\in V^*$ 的意义，并举例说明.

11. 试证 $\forall\sigma\in\mathfrak{S}r)$，$\sigma:T^r(V)\to T^r(V)$ 是自同态映射.

12. 试证对称张量在对称算子作用下不变，反对称张量在反对称算子作用下不变.

13. 为何在证明反对称性时取的是置换 $\tau=\begin{pmatrix}1 & \cdots & k & k+1 & \cdots & k+l\\ 1+l & \cdots & k+l & 1 & \cdots & l\end{pmatrix}$?

14. 试证定理 6.2.7、定理 6.2.8.

15. 试证定理 6.2.9；试证定理 6.2.12、定理 6.2.13.

16. 试证(6.2.24)式与(6.2.25)式分别为空间 $\Lambda(V)$ 与 $\Lambda(V^*)$ 的基.

17. 对于定理 6.3.2，给出对于四维情形的证明.

18. 对于外微分式空间，在 $n=4$ 时，给出具体的表示式.

19. 试证 p 次外微分式 ω 与 q 次外微分式 θ 的“反莱布尼茨公式”

$$\mathrm{d}(\omega ? \theta)=\mathrm{d}\omega ? \theta+(-1)^p\omega ? \mathrm{d}\theta.$$

20. 求 $\omega=(y^3-z^3)\mathrm{d}x+(z^3-x^3)\mathrm{d}y+(x^3-y^3)\mathrm{d}z$ 的 1 次、2 次外微分式.

21. 设 $\omega=(x^2+y^2+z^2)\mathrm{d}x+(2xy-z^2)\mathrm{d}y+(xy+yz+zx)\mathrm{d}z$，求 $\mathrm{d}\omega$.

22. 对于外微分式空间 $A(\mathbb{R}^3)$，设 2 次外微分式 $\omega=2x\mathrm{d}y?\mathrm{d}z+2y\mathrm{d}z?\mathrm{d}x+3z\mathrm{d}x?\mathrm{d}y$，求外微分 $\mathrm{d}\omega$.

23. 设 V 是二维线性空间，若基为$\{e_1,e_2\}$，求反对称 1 次协变张量的表示；若 $\omega\in A(M)$，求 $\mathrm{d}\omega$.

24. 设 $\omega=-y\mathrm{d}x+x\mathrm{d}y$ 是定义在$\mathbb{R}^3$ 上的 1 次外微分式，$\alpha=\omega|_{S^2}$ 是 ω 在$\mathbb{R}^3$ 的单位球面 $S^2=\{(x,y,z)\in\mathbb{R}^3:x^2+y^2+z^2=1\}$上的限制，试求在 S^2 的局部坐标系$\{(U;u,v)\}$上的$\mathrm{d}(\alpha|_U)$与 $\mathrm{d}(\alpha|_V)$，其中 $U=\{(x,y,z)\in S^2:x=u,y=v,z=\sqrt{1-u^2-v^2}\}$.

25. 设 $\omega=xy\mathrm{d}x+z\mathrm{d}y-yz\mathrm{d}z$，光滑映射 $f:\mathbb{R}^2\to\mathbb{R}^3$ 由 $f(u,v)=(uv,u^2,3u+v)$给出，求 $f^*\omega$，$f^*\mathrm{d}\omega$.

26. 给出定理 6.4.1 与定理 6.4.2 的证明.

27. 写出(6.4.5)式在 $n=4$ 情形下的表示.

28. 在 $n=11$ 时，解释 Stokes 公式.

29. 设 $T^2=\{(x,y,z)\in\mathbb{R}^3:(\sqrt{x^2+y^2}-R)^2+z^2=r^2\}(0<r<R)$是$\mathbb{R}^3$ 中的环面，设$I:T^2\to\mathbb{R}^3$ 是嵌入映射，并设 T^2 的正向为外法线方向，若 $\omega=x\mathrm{d}y?\mathrm{d}z+y\mathrm{d}z?\mathrm{d}x+z\mathrm{d}x?\mathrm{d}y$，试求 $\int_{T^2}I^*\omega$.

30. 将 Riemann 几何的一般理论化到 $n=4$，$n=11$ 的情形，并写出具体的表示式.

31. 设$\mathbb{R}^{n+1}$中的单位球面为 $S^n(1)=\{(x_1,\cdots,x_n,x_{n+1})\in\mathbb{R}^{n+1}:x_1^2+\cdots+x_n^2+x_{n+1}^2=1\}$，试证 $S^n(1)$是$\mathbb{R}^{n+1}$中的 n 维 Riemann 流形.

32. 求$\mathbb{R}^3$ 中的下列曲面的第一基本形式：

(1) $r=(u\cos v,u\sin v,av)$，$a>0$；

(2) $r=(\cosh u\cos v,\cosh u\sin v,u)$.

33. 以 $n=5$ 与 $n=6$ 为具体背景，熟悉 Riemann 几何的一般理论.

第7章 补充知识

作为补充知识,本章介绍几个经常遇到的内容,请读者自己阅读.

7.1 变分方法

变分法是处理极大、极小问题的一种数学方法. 有限个自变量的极值问题通常在高等数学中进行讨论,而变分法则是处理泛函的极值问题.

7.1.1 变分与变分问题

1. 例子

首先给出一些例子.

例 7.1.1 二次整式的极值问题

考虑 n 个实变量 $u_1,\cdots,u_n\in\mathbb{R}$ 的二次整式

$$J(u)=\sum_{j,k=1}^{n}a_{jk}u_ju_k+2\sum_{j=1}^{n}a_ju_j+a_0, \tag{7.1.1}$$

其中系数 a_{jk},a_j,a_0 是实数,n 阶方阵$[a_{jk}]$为实对称的,$a_{jk}=a_{kj}$, $j,k=1,2,\cdots,n$.

为求使得 $J(u)$取极值的 u,依高等数学中的方法,先求其“静止点”. 亦即,求解

$$\frac{1}{2}\frac{\partial J}{\partial u_j}=\sum_{k=1}^{n}a_{jk}u_k+a_j=0,\quad j=1,2,\cdots,n. \tag{7.1.2}$$

令 $A=[a_{jk}],A^{-1}=[a_{jk}]^{-1}=[a'_{jk}],b=\begin{bmatrix}a_1\\a_2\\\vdots\\a_n\end{bmatrix},u=\begin{bmatrix}u_1\\u_2\\\vdots\\u_n\end{bmatrix}$,则(7.1.2)式中的方程化为 $Au+b=0$. 解此方程,得

$$\tilde{u} = -A^{-1}b = -[a'_{jk}]\begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}, \tag{7.1.3}$$

$\tilde{u}$是 $J(u)$的**静止点**，$J(\tilde{u})$是其静止值. 我们有

$$\begin{aligned} J(\tilde{u}) &= (\tilde{u})^{\mathrm{T}} A \tilde{u} + 2b^{\mathrm{T}} \tilde{u} + a_0 = \{-b^{\mathrm{T}}(A^{-1})^{\mathrm{T}}\} A\{-A^{-1}b\} + 2b^{\mathrm{T}}\{-A^{-1}b\} + a_0 \\ &= b^{\mathrm{T}}\{(A^{-1})^{\mathrm{T}}\} AA^{-1}b - 2b^{\mathrm{T}}A^{-1}b + a_0 = b^{\mathrm{T}}A^{-1}b - 2b^{\mathrm{T}}A^{-1}b + a_0 \\ &= -b^{\mathrm{T}}A^{-1}b + a_0 = -\sum_{j,k=1}^{n} a'_{jk}a_j a_k + a_0, \end{aligned}$$

这里需要假设 A^{-1}的对称性. 令 $u_j - \tilde{u}_j = \delta\tilde{u}_j\,(j=1,\cdots,n)$为$\tilde{u}_j$ 的增量，并利用

$$\delta\tilde{u}_j \delta\tilde{u}_k = (u_j - \tilde{u}_j)(u_k - \tilde{u}_k) = u_j u_k - \tilde{u}_j u_k - u_j\tilde{u}_k + \tilde{u}_j\tilde{u}_k,$$

则

$$\begin{aligned} J(u) - J(\tilde{u}) &= u^{\mathrm{T}}Au + 2b^{\mathrm{T}}u + a_0 - (\tilde{u})^{\mathrm{T}}A\tilde{u} - 2b^{\mathrm{T}}\tilde{u} - a_0 \\ &= u^{\mathrm{T}}Au + 2b^{\mathrm{T}}u - (\tilde{u})^{\mathrm{T}}A\tilde{u} - 2b^{\mathrm{T}}\tilde{u} \\ &= \sum_{j,k=1}^{n} a_{jk}(u_j u_k - \tilde{u}_j\tilde{u}_k) + 2\sum_{j=1}^{n} a_j(u_j - \tilde{u}_j) = \sum_{j,k=1}^{n} a_{jk}\delta\tilde{u}_j\delta\tilde{u}_k. \end{aligned}$$

因此，问题归结为方阵 $A=[a_{jk}]$与二次型 $x^{\mathrm{T}}Ax = \sum_{j,k=1}^{n} a_{jk}x_j x_k$，$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$的讨论. 由微积分的知识知

$J(\tilde{u})$ 为极小值 $\Leftrightarrow$ 二次型$(\delta\tilde{u})^{\mathrm{T}}A(\delta\tilde{u})$ 为正定的；

$J(\tilde{u})$ 为极大值 $\Leftrightarrow$ 二次型$(\delta\tilde{u})^{\mathrm{T}}A(\delta\tilde{u})$ 为负定的.

例 7.1.2 特征值问题

考虑 n 个实变量 $u_1,\cdots,u_n\in\mathbb{R}$ 的商式

$$J(u) = \frac{\sum_{j,k=1}^{n} a_{jk}u_j u_k}{\sum_{j,k=1}^{n} b_{jk}u_j u_k} = \frac{u^{\mathrm{T}}Au}{u^{\mathrm{T}}Bu} = \frac{(u,Au)}{(u,Bu)}, \tag{7.1.4}$$

其中系数 a_{jk}，b_{jk}是实数，方阵 $A=[a_{jk}]$，$B=[b_{jk}]$为实对称的，$a_{jk}=a_{kj}$，$b_{jk}=b_{kj}$，且$|B|\neq 0$. 求(7.1.4)式的极值与极值点.

由于 $J(u)=\dfrac{u^{\mathrm{T}}Au}{u^{\mathrm{T}}Bu}=\dfrac{(u,Au)}{(u,Bu)}$是零次齐次的，亦即 $J(\lambda u)=J(u)$，故可令"内积"$(u,Bu)=1$. 此问题有几何应用：求椭圆体的最小(最大)半径.

为求 $J(u)$的极小值，写出"静止条件"

$$\frac{\partial J}{\partial u_j}=0,\quad j=1,2,\cdots,n. \tag{7.1.5}$$

将(7.1.4)写为 $(u^{\mathrm{T}}Bu)J(u)=u^{\mathrm{T}}Au$,并令 $\lambda=J(u)$,于是,静止条件化为

$$\begin{cases}\displaystyle\sum_{j,k=1}^{n}(a_{jk}-\lambda b_{jk})u_k=0,\\ \lambda=J(u),\end{cases}$$

或 $Au=\lambda Bu$. 演化为线性代数中的问题,就是求矩阵 A 的特征值(固有值)问题,亦即,求特征方程 $|A-\lambda I|=0$ 的解,即求特征值 $\lambda_1,\lambda_2,\cdots,\lambda_n$,而对应于特征值 λ 的 u,满足 $J(\tilde{u})=\lambda$,称为特征向量(固有向量).

这个例子的更一般的情形是函数的极值问题,即求

$$J(u)=J(u_1,\cdots,u_n),\quad (u_1,\cdots,u_n)\in D\subset\mathbb{R}^n$$

的极值. 回顾极小点与静止点的定义:

极小点 $\tilde{u}\in D$,使得 $J(u)\geqslant J(\tilde{u})$(广义极小);$\tilde{u}\in D$,使得 $J(u)>J(\tilde{u})$(狭义极小).

为求极值点,可以从静止点中寻找.

静止点 ① $\tilde{u}\in D,\tilde{u}\notin\partial D$,使得 $\dfrac{\partial J}{\partial u_j}=0,j=1,2,\cdots,n$; ② $\tilde{u}\in\partial D$.

然而,为确定静止点是否是极值点,还需要用充分条件来判别.

例 7.1.3 捷线问题

考虑一质点沿着固定的光滑曲线,仅借助重力的作用,从 P_0 滑到 P_1,为使滑行时间最短,问需通过怎样的曲线?

设这个曲线的方程为 $u=u(x)$,重力加速度为 g,曲线上的弧微分为

$$\mathrm{d}s=\sqrt{1+[u'(x)]^2}\,\mathrm{d}x.$$

设点 $P(x,u)$是曲线上的任一点,如图 7.1.1 所示,该点的瞬时速度为 $\sqrt{2gu}$,而质点通过弧 $\mathrm{d}s$ 的时间是

$$\mathrm{d}t=v^{-1}\mathrm{d}s=\frac{1}{\sqrt{2gu}}\sqrt{1+[u'(x)]^2}\,\mathrm{d}x,$$

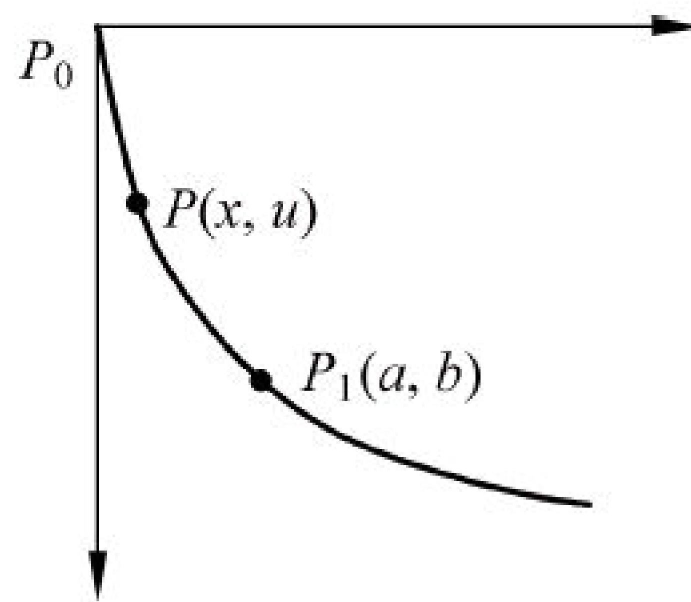


图 7.1.1 捷线

不失一般性,令 $2g=1$,则从 P_0 到 P_1 的时间为 $J(u)=\displaystyle\int_0^a\frac{\sqrt{1+[u'(x)]^2}}{\sqrt{u(x)}}\mathrm{d}x$,于是,问题是求使得 $J(u)$ 成为最小的、满足 $u(0)=0,u(a)=b$ 的函数 $u(x)$,亦即,求问题

$$\begin{cases}J(u)=\displaystyle\int_0^a\frac{\sqrt{1+[u'(x)]^2}}{\sqrt{u(x)}}\mathrm{d}x,\\ u(0)=0,\quad u(a)=b\end{cases} \tag{7.1.6}$$

的最小函数 $u(x)$.

2. Euler 方程

考虑

$$\begin{cases} J(u) = \int_{a_0}^{a_1} F(x,u,u')\mathrm{d}x, \\ u(a_0) = b_0,\ u(a_1) = b_1, \end{cases} \tag{7.1.7}$$

求满足问题(7.1.7)的最小函数 $u=\tilde{u}(x)$,如图 7.1.2 所示.

解决的办法是:取 $\eta(x)$,满足 $\eta(a_0)=\eta(a_1)=0$,又 $\forall\varepsilon\in\mathbb{R}$,令

$$u_\varepsilon(x) = \tilde{u}(x) + \varepsilon\eta(x). \tag{7.1.8}$$

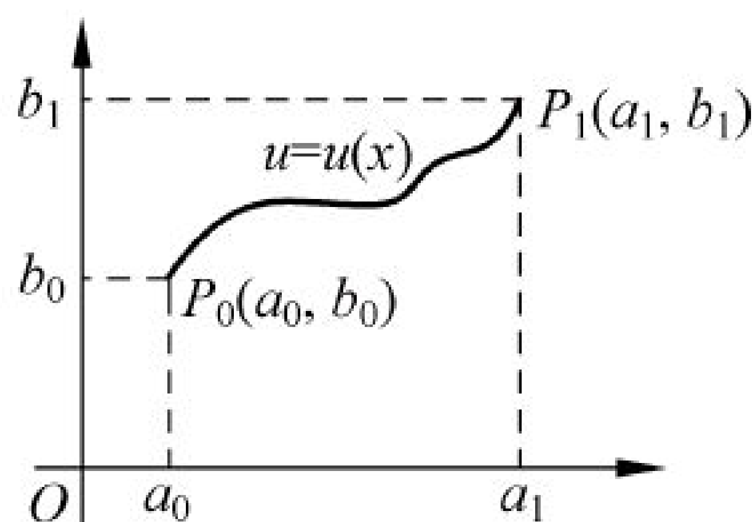


图 7.1.2 $u=u(x)$

显然,

$$\begin{cases} u_\varepsilon(a_0) = \tilde{u}(a_0) = b_0, \\ u_\varepsilon(a_1) = \tilde{u}(a_1) = b_1. \end{cases}$$

因此,$u_\varepsilon(x)$可视为满足问题(7.1.7)的"可取函数",将其称为由$\tilde{u}(x)$变分所得的函数. 代入 $J(u)$得

$$J(u_\varepsilon) = \int_{a_0}^{a_1} F(x,u_\varepsilon,u'_\varepsilon)\mathrm{d}x.$$

$J(u_\varepsilon)$是 ε 的函数,当 $\varepsilon=0$ 时,$u_0(x)=\tilde{u}(x)$使 $J(u_\varepsilon)$取最小值,于是

$$\begin{aligned}
\frac{\partial}{\partial\varepsilon}J(u_\varepsilon)\Big|_{\varepsilon=0} &= \frac{\partial}{\partial\varepsilon}\int_{a_0}^{a_1} F(x,u_\varepsilon,u'_\varepsilon)\mathrm{d}x\Big|_{\varepsilon=0} \\
&= \frac{\partial}{\partial\varepsilon}\int_{a_0}^{a_1} F(x,\tilde{u}(x)+\varepsilon\eta(x),(\tilde{u})'(x)+\varepsilon\eta'(x))\mathrm{d}x\Big|_{\varepsilon=0} \\
&= \int_{a_0}^{a_1}\left\{\frac{\partial}{\partial u}F(x,\tilde{u}(x),(\tilde{u})'(x))\eta(x) + \frac{\partial}{\partial u'}F(x,\tilde{u}(x),(\tilde{u})'(x))\eta'(x)\right\}\mathrm{d}x \\
&= \int_{a_0}^{a_1}\widetilde{F}_u(x)\eta(x)\mathrm{d}x + \widetilde{F}_{u'}(x)\eta(x)\Big|_{a_0}^{a_1} - \int_{a_0}^{a_1}\left(\frac{\mathrm{d}}{\mathrm{d}x}\widetilde{F}_{u'}(x)\right)\eta(x)\mathrm{d}x \\
&= \int_{a_0}^{a_1}\left(\widetilde{F}_u - \frac{\mathrm{d}\widetilde{F}_{u'}}{\mathrm{d}x}\right)\eta(x)\mathrm{d}x = 0.
\end{aligned}$$

这样,$u=\tilde{u}(x)$必须满足微分方程

$$\widetilde{F}_u - \frac{\mathrm{d}\widetilde{F}_{u'}}{\mathrm{d}x} = 0, \tag{7.1.9}$$

详细地写出,即为

$$\widetilde{F}_u - \frac{\mathrm{d}\widetilde{F}_{u'}}{\mathrm{d}x} = F_u(x,u,u'(x)) - \frac{\mathrm{d}}{\mathrm{d}x}F_{u'}(x,u,u'(x)) = 0. \tag{7.1.10}$$

方程(7.1.9)与方程(7.1.10)称为 **Euler 方程**,它是使得 $J(u)$取得极值的必要条件.

这里我们用到了**变分基本引理**:若连续函数 $f(x)$,对于满足边界条件 $\eta(a_0)=\eta(a_1)=$

0 的、具有适当光滑性的任意函数 $\eta(x)$ 都满足$\int_{a_0}^{a_1} f(x)\eta(x)\mathrm{d}x = 0$,则 $f(x) \equiv 0$.

将在下节中重新回到这个引理.

3. Euler 方程积分法

Euler 方程(7.1.10)是一个微分方程,我们只能在一些特殊情况下进行积分.

1) 当 F 不含 u 时,$F(x,u')=0$:方程(7.1.10)成为$\dfrac{\mathrm{d}F_{u'}}{\mathrm{d}x}=0$,得到一个首次积分 $F_{u'}(x,u'(x))=C$.

2) 当 F 不含 x 时,$F(u,u')=0$:方程(7.1.10)成为

$$\frac{\mathrm{d}}{\mathrm{d}x}(F - u'F_{u'}) = 0, \tag{7.1.11}$$

这是因为

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}(F - u'F_{u'}) &= F_x + F_u \cdot u' + F_{u'} \cdot u'' - u'' \cdot F_{u'} - u' \frac{\mathrm{d}}{\mathrm{d}x}F_{u'} \\ &= F_u \cdot u' - u' \frac{\mathrm{d}}{\mathrm{d}x}F_{u'} = u'\left(F_u - \frac{\mathrm{d}}{\mathrm{d}x}F_{u'}\right) = 0.\end{aligned}$$

解方程(7.1.11),得首次积分 $F-u'F_{u'}=C$.

例 7.1.4 推广的捷线问题:$J(u)=\int_{a_0}^{a_1} f(u)\sqrt{1+[u'(x)]^2}\,\mathrm{d}x$

对于 $J(u)$,有 $F(x,u,u')=f(u)\sqrt{1+[u'(x)]^2}$. 因 F 不含 x,故由

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}(F - u'F_{u'}) &= \frac{\mathrm{d}}{\mathrm{d}x}\left[f(u)\sqrt{1+(u')^2} - u'f(u)\frac{u'}{\sqrt{1+(u')^2}}\right] \\ &= \frac{\mathrm{d}}{\mathrm{d}x}\left[\frac{f(u)}{\sqrt{1+(u')^2}}\right] = 0,\end{aligned}$$

得到一个首次积分 $\dfrac{f(u)}{\sqrt{1+(u')^2}}=c$. 由此,得$\dfrac{c^2(u')^2}{f^2(u)-c^2}=1$. 于是,使得极值问题

$$J(u) = \int_{a_0}^{a_1} f(u)\sqrt{1+[u'(x)]^2}\,\mathrm{d}x$$

取极小值的 $u(x)$ 满足

$$c\int [f^2(u) - c^2]^{-\frac{1}{2}}\mathrm{d}u = \pm(x-\xi). \tag{7.1.12}$$

另一方面,u 取常数也可能是问题的解.

例 7.1.5 捷线问题的解

对于(7.1.6)式$\begin{cases} J(u) = \displaystyle\int_0^a \frac{\sqrt{1+[u'(x)]^2}}{\sqrt{u(x)}}\mathrm{d}x, \\ u(0)=0,\ u(a)=b, \end{cases}$可利用例 7.1.4 的结果,因为 $f(u)=\dfrac{1}{\sqrt{u}}$,

代入例 7.1.4 中的公式,得到 $c\int\left[\frac{u}{1-c^2u}\right]^{\frac{1}{2}}\mathrm{d}u=\pm(x-\xi)$. 令 $u=c^{-2}\sin^2\theta$,于是,问题的解为

$$\begin{cases} x=\frac{1}{2c^2}(2\theta-\sin 2\theta)+\xi, \\ u=\frac{1}{2c^2}(1-\cos 2\theta), \end{cases} \tag{7.1.13}$$

这是一条摆线,它是一个以 $\frac{1}{2c^2}$ 为半径的圆,沿 Ox 轴转动时,圆周上一点所描出的曲线的轨迹.

由于没有满足 $f'(b)=0$ 的 b,所以 $u=b$ 不是极值曲线. 又虽然 $u=c^{-2}$ 也是函数方程 $F-u'F_{u'}=C$ 的奇异点,但是却不是 Euler 方程的解,因此,只有(7.1.13)式才是极值曲线.

为了使得所求的解满足边界条件,例如,采取 $\theta=0$ 作为出发点,就有 $\xi=0$. 常数 c^2 可由曲线通过(a_1,b_1)的条件来确定.

图 7.1.4 中,圆心为 A,摆线上的一点为 B,与 Ox 的切点为 C,$\angle BAC=2\theta$,$AC=\frac{1}{2c^2}$.

由几何直观,知 c^2 存在,且是惟一确定的. 若$\frac{b_1}{a_1}\geqslant\frac{2}{\pi}$,则质点单调地沿着摆线落到地面;若 $\frac{b_1}{a_1}<\frac{2}{\pi}$,则质点落下一段时间后又上升,然后才达到目的地.

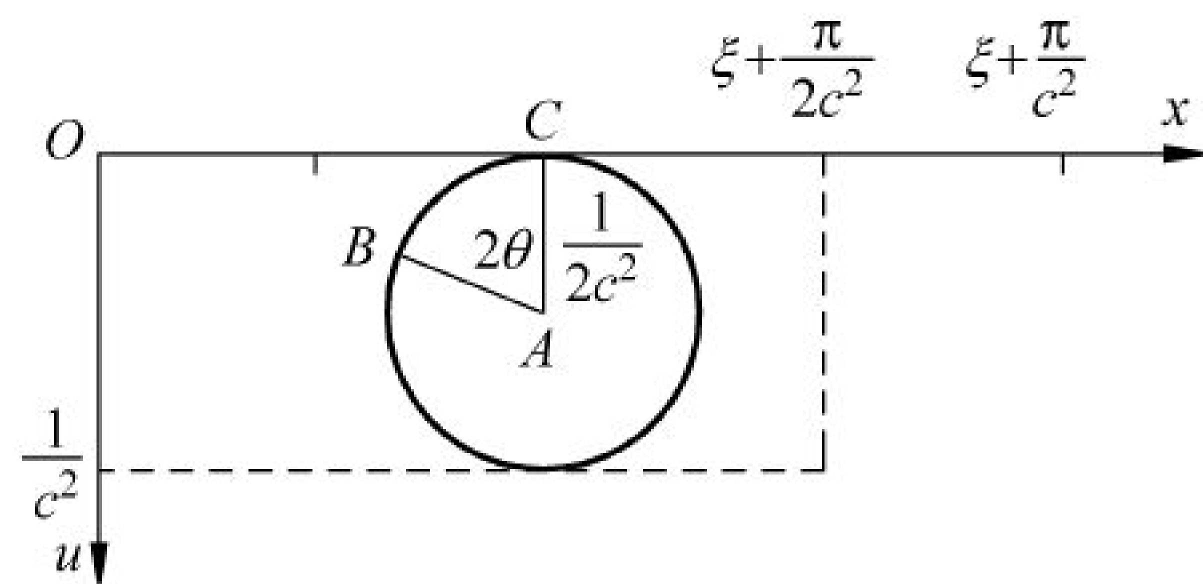


图 7.1.4 摆线

更一般地,当给定了任意两点(a_0,b_0),(a_1,b_1),$b_0\geqslant 0$ 与 $b_1\geqslant 0$ 时,通过这两点的极小值曲线必存在,且是惟一确定的.

例 7.1.6 另一些极值问题

在例 7.1.4 中,取 $f(u)=\sqrt{u}$,是力学中的最小作用原理在抛射体运动中的应用(仍取铅垂向下方向为 u 轴). 于是,由 $J(u)=\int_{a_0}^{a_1}\sqrt{u(x)}\sqrt{1+[u'(x)]^2}\,\mathrm{d}x$,得一个首次积分(7.1.12),有

$$u=\frac{1}{4c^2}(x-\xi)^2+c^2. \tag{7.1.14}$$

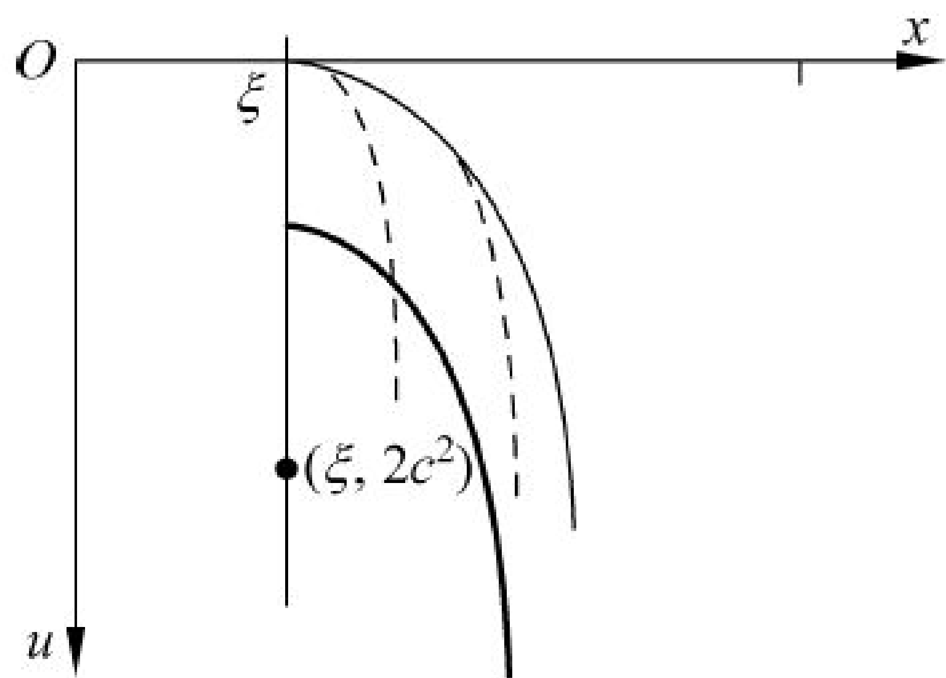


图 7.1.5 抛物线

根据与例 7.1.5 同样的讨论,可知(7.1.14)式是本问题的惟一的极值曲线,它是以平行于 u 轴的直线 $x=\xi$ 为轴、以实轴 $u=0$ 为准线的抛物线,焦点在$(\xi,2c^2)$.

参数(积分常数)ξ,c 用来确定边界条件,事实上,所有极值曲线有一条通过给定的点(a_0,b_0)的抛物线(是包络)

$$u = \frac{(x-a_0)^2}{4b_0}.$$

例 7.1.7 在例 7.1.4 中,取 $f(u)=u$,是使得通过给定两点 $P_0(a_0,b_0)$,$P_1(a_1,b_1)$的曲线 $u=u(x)$绕 Ox 轴旋转所生成的曲面面积为最小的问题,即 $J(u)=\int_{a_0}^{a_1} u(x)\sqrt{1+[u'(x)]^2}\,\mathrm{d}x$. 由于 F 不显含 x,故按上面的解法,得到

$$u = c\cosh\frac{x-\xi}{c},$$

这是悬链线(eatenary),当初始条件(a_0,b_0),(a_1,b_1)在横轴方向相距太远时,则不存在极值曲线;若相距较靠近,则会有两条极值嫌疑曲线.

7.1.2 变分原理

1. 变分基本引理

前面我们用到了变分原理,现在给出证明.

变分基本引理 若连续函数 $f(x)$,对于满足边界条件 $\eta(a_0)=\eta(a_1)=0$ 的、具有适当光滑性的任意函数 $\eta(x)$ 都满足 $\int_{a_0}^{a_1} f(x)\eta(x)\mathrm{d}x=0$,则 $f(x)\equiv 0$.

证 若不然,$\exists x_0\in(a_0,a_1)$,使得 $f(x_0)\neq 0$,例如 $f(x_0)>0$,则由 $f(x)$的连续性,必有 $\zeta>0$,使得在$(x_0-\zeta,x_0+\zeta)$上,有 $f(x)>0$. 于是,必可取光滑函数 $\eta(x)$,满足 $\eta(a_0)=\eta(a_1)=0$、在$(x_0-\zeta,x_0+\zeta)$上有 $\eta(x)>0$、在$(x_0-\zeta,x_0+\zeta)$外有 $\eta(x)=0$,于是

$$\int_{a_0}^{a_1} f(x)\eta(x)\mathrm{d}x = \int_{x_0-\zeta}^{x_0+\zeta} f(x)\eta(x)\mathrm{d}x > 0,$$

与假设矛盾.

2. 最小值、极小值

对于前面讨论的(7.1.7) 极值问题 $\begin{cases} J(u)=\int_{a_0}^{a_1} F(x,u,u')\mathrm{d}x, \\ u(a_0)=b_0,\ u(a_1)=b_1, \end{cases}$ 其中 $F(x,u,u')$ 是区域 $D\subset\mathbb{R}^3$ 上具有充分光滑性的函数,而泛函

$$J(u) = \int_{a_0}^{a_1} F(x,u,u')\mathrm{d}x$$

中的函数 $u=u(x)$是逐段光滑的,满足(7.1.7)式的 $u=u(x)$称为“**可取函数**”.

1) 泛函 $J(u)$的最小值

若所有可取函数 $u=u(x)$中,存在 $u=\tilde{u}(x)$,使得

$$J(\tilde{u}) \leqslant J(u) \tag{7.1.15}$$

对于任意可取函数 $u=u(x)$都成立,则称 $J(\tilde{u})$为 $J(u)$最小值,而称$\tilde{u}$为最小函数. 若当 $u\neq$

$\tilde{u}$时，成立 $J(\tilde{u})<J(u)$，则称 $J(\tilde{u})$ 为 $J(u)$ 的**狭义最小值**，否则，称 $J(\tilde{u})$ 为 $J(u)$ 的**广义最小值**.

2）泛函 $J(u)$ 的极小值

① 强极小值

若取适当的正数 $\delta>0$，满足

$$|u(x)-\tilde{u}(x)|<\delta,\quad a_0\leqslant x\leqslant a_1 \tag{7.1.16}$$

的可取函数 u 的全体，称为$\tilde{u}$的**近旁**. 于是，若存在适当的 $\delta>0$，使得满足(7.1.16)式的可取函数 u，都成立 $J(\tilde{u})\leqslant J(u)$，则称 $J(\tilde{u})$ 为 $J(u)$ 的**强极小值**；也可分为**狭义强极小值**与**广义强极小值**.

② 弱极小值

取适当的正数 $\delta>0$，考虑满足

$$|u(x)-\tilde{u}(x)|<\delta,\ |u'(x)-(\tilde{u})'(x)|<\delta,\quad a_0\leqslant x\leqslant a_1 \tag{7.1.17}$$

的可取函数 u 的全体；若存在适当的 $\delta>0$，使得满足(7.1.17)式的可取函数 u，都成立 $J(\tilde{u})\leqslant J(u)$，则称 $J(\tilde{u})$ 为 $J(u)$ 的**弱极小值**；也可分为**狭义弱极小值**与**广义弱极小值**.

最大值，极大值可类似定义.

7.1.3 更一般的变分问题

1. 多个函数的情形

考虑多个未知函数 $u_1(x),\cdots,u_k(x)$ 的情形

$$\begin{cases} J(u)=\int_{a_0}^{a_1}F(x,u_1,\cdots,u_k,u_1',\cdots,u_k')\mathrm{d}x, \\ u_j(a_0)=b_{0j},\ u_j(a_1)=b_{1j},\quad 1\leqslant j\leqslant k, \end{cases}$$

得到的 Euler 方程组

$$F_{u_j}-\frac{\mathrm{d}F_{u_j'}}{\mathrm{d}x}=0,\quad j=1,\cdots,k, \tag{7.1.18}$$

并且满足正则条件 $\det\left[\frac{\partial^2 F}{\partial u_j'\partial u_s'}\right]_{1\leqslant j,s\leqslant k}\neq 0$.

在分析力学中，若时间取作 x，取力学系统的一般坐标为 $u_1(x),\cdots,u_k(x)$，取表达动能与位能之差的 Lagrange 函数作为 F，则由 J 作为极值的条件，这个系统的运动便确定了，而 Euler 方程组(7.1.18)就是相应的运动方程.

2. 含高阶导函数的情形

考虑含未知函数 $u(x)$，$x\in[a_0,a_1]$，及其各高阶导数 $u'(x),\cdots,u^{(k)}(x)$，$x\in[a_0,a_1]$，的情形，$k\in\mathbb{Z}^+$：

$$\begin{cases} J(u)=\int_{a_0}^{a_1}F(x,u,u',\cdots,u^{(k)})\mathrm{d}x, \\ u(a_0)=b_0,\cdots,u^{(k-1)}(a_0)=b_0^{(k-1)}, \\ u(a_1)=b_1,\cdots,u^{(k-1)}(a_1)=b_1^{(k-1)}, \end{cases}$$

得到的 Euler 方程

$$F_u-\frac{\mathrm{d}}{\mathrm{d}x}F_{u'}+\frac{\mathrm{d}^2}{\mathrm{d}x^2}F_{u''}-\cdots+(-1)^k\frac{\mathrm{d}^k}{\mathrm{d}x^k}F_{u^{(k)}}=0, \tag{7.1.19}$$

方程(7.1.19)一般是 $2k$ 阶微分方程.

3. 含两个以上自变量的情形

考虑最简单的含两个自变量 x,y,一个未知函数 $u(x,y)$及其一阶偏导数的情形:

$$\begin{cases} J(u)=\iint\limits_G F(x,y,u,u_x,u_y)\mathrm{d}x\mathrm{d}y, \\ u\big|_{\partial G}=f(s),\ s\in\partial G, \end{cases}$$

得到的 Euler 方程

$$F_u-\frac{\partial}{\partial x}F_{u_x}-\frac{\partial}{\partial y}F_{u_y}=0, \tag{7.1.20}$$

方程(7.1.20)一般是关于 u 的二阶偏微分方程.

4. 等周问题(isoperimetric problem)

等周问题是由 $J(u)=\int_{a_0}^{a_1}F(x,u,u')\mathrm{d}u$ 在 $K(u)=\int_{a_0}^{a_1}G(x,u,u')\mathrm{d}x$ 取固定值时取最小(极小)值的问题,亦即在“周长”一定的所有闭曲线中,求其所围面积为最大的一条曲线.

设 $u=\tilde{u}$为所求的极小函数,考虑变分

$$u=\tilde{u}(x)+\varepsilon_1\eta_1(x)+\varepsilon_2\eta_2(x),$$

代入 $J(u)=\int_{a_0}^{a_1}F(x,u,u')\mathrm{d}u$ 与 $K(u)=\int_{a_0}^{a_1}G(x,u,u')\mathrm{d}x$,得到两个变量$\varepsilon_1,\varepsilon_2$ 的函数 $J(\varepsilon_1,\varepsilon_2)$,$K(\varepsilon_1,\varepsilon_2)$. 在假设 $K(\varepsilon_1,\varepsilon_2)=c$ 之下,当 $\varepsilon_1=\varepsilon_2=0$ 时,$J(u)$ 取最小值,故由待定乘数法,得到

$$\left(\frac{\partial J}{\partial\varepsilon_1}\right)\bigg|_{\varepsilon_1=\varepsilon_2=0}-\lambda\left(\frac{\partial K}{\partial\varepsilon_1}\right)\bigg|_{\varepsilon_1=\varepsilon_2=0}=0,\quad \left(\frac{\partial J}{\partial\varepsilon_2}\right)\bigg|_{\varepsilon_1=\varepsilon_2=0}-\lambda\left(\frac{\partial K}{\partial\varepsilon_2}\right)\bigg|_{\varepsilon_1=\varepsilon_2=0}=0,$$

由此导出 Euler 方程

$$(F-\lambda G)_u-\frac{\mathrm{d}}{\mathrm{d}x}(F-\lambda G)_{u'}=0, \tag{7.1.21}$$

方程(7.1.21)与以前导出的 Euler 方程类似.

例 7.1.8 等周问题的例子

对于例 7.1.7 中的最小值问题 $J(u)=\int_{a_0}^{a_1}u(x)\sqrt{1+[u'(x)]^2}\,\mathrm{d}x$，加上条件 $K(u)=\int_{a_0}^{a_1}\sqrt{1+(u')^2}\,\mathrm{d}x=l$，就成为一个等周问题.

在方程(7.1.12)中，令 $f(u)=u-\lambda$，便可得到 Euler 方程的积分，并且不显含 x，故按上面的解法，得到

$$u=\lambda+c\cosh\frac{x-\xi}{c}, \tag{7.1.22}$$

此即极小曲线的表示式. 再将其代入 $K(u)=\int_{a_0}^{a_1}\sqrt{1+(u')^2}\,\mathrm{d}x=l$，得到

$$\sinh\frac{a_1-\xi}{c}-\sinh\frac{a_0-\xi}{c}=\frac{1}{c}. \tag{7.1.23}$$

曲线通过两点 $P_0(a_0,b_0)$，$P_1(a_1,b_1)$ 的条件是

$$\lambda+c\cosh\frac{a_0-\xi}{c}=b_0,\quad \lambda+c\cosh\frac{a_1-\xi}{c}=b_1. \tag{7.1.24}$$

由以上三式可决定常数 ξ,c,λ. 易见，当 $l^2\geqslant(a_1-a_0)^2+(b_1-b_0)^2$ 时，必定存在惟一的解(此时，曲线的长度大于两点间的距离).

7.2 Banach 空间中的几个重要定理

在高等数学(参看[11])的第 6 章中，简单介绍了 Banach 空间中的微分学，对于两个 Banach 空间 X,Y 之间的映射 $f:X\to Y$，定义并研究了极限、导数、级数等. 关于积分，考虑定义在实轴的区间 $[a,b]$ 上取值于 Banach 空间 Y 的映射 $f:[a,b]\to Y$ 的积分，其本质与 Riemann 积分相同，读者可参考[10]. 本节介绍几个重要定理：Stone-Weierstrass 定理、隐映射定理、逆映射定理以及不动点原理.

7.2.1 Stone-Weierstrass 定理

首先回顾经典的 Weierstrass 定理.

定理 7.2.1 设 $f(x)$ 是定义在区间 $[a,b]$ 上的实值连续函数，则存在实系数多项式序列 $\{p_n(x)\}$，使得 $p_n(x)$ 在 $[a,b]$ 上一致收敛于 $f(x)$.

证明可参看[6].

为介绍 Stone-Weierstrass 定理，给出几个定义.

定义 7.2.1(集合 A 区分 X 的点) 设 X 为任一集合，$\mathfrak{B}(X)$ 是 X 上实值有界函数的全体，即

$$\mathfrak{B}(X)=\{f:X\to\mathbb{R},f\text{ 是 }X\text{ 上的有界函数}\},$$

$A\subset\mathfrak{B}(X)$为$\mathfrak{B}(X)$的一个子集．称**集合 A 区分 X 的点**，若对于 X 中的任意两个不同的元 $x,y\in X,x\neq y$，存在 $f\in A$，使得 $f(x)\neq f(y)$．

在下面的定理中，我们用到的是 X 为 T_2 型紧致拓扑空间，因此 X 上的实值连续函数空间 $C_{\mathbb{R}}(X)$是$\mathfrak{B}(X)$的子空间，即 $C_{\mathbb{R}}(X)\subset\mathfrak{B}(X)$．

$(C_{\mathbb{R}}(X),+,\alpha\cdot,\|f\|_{C_{\mathbb{R}}(X)},f\cdot g)$在加法、数乘、范数、乘法之下成为一个具有(乘法)单位元的 Banach 代数，$e\equiv e_{C_{\mathbb{R}}(X)}=1$．

定理 7.2.2(Stone-Weierstrass 定理) 设 X 为 T_2 型紧致拓扑空间，$C_{\mathbb{R}}(X)$为 X 上的 Banach 代数．若子集 $A\subset C_{\mathbb{R}}(X)$是 $C_{\mathbb{R}}(X)$的一个子代数，且含有 $C_{\mathbb{R}}(X)$的乘法单位元 e，则 $\overline{A}=C_{\mathbb{R}}(X)$，当且仅当 A 区分 X 的点．

证 必要性 若 $\overline{A}=C_{\mathbb{R}}(X)$，证明 A 区分 X 的点．

因为 X 是 T_2 型空间，因此是 T_4 型的，故任取 $x,y\in X,x\neq y$，据 Urysohn 引理，存在连续函数 $f\in C_{\mathbb{R}}(X)$，使得

$$f(x)=0,\quad f(y)=1. \tag{7.2.1}$$

由必要条件 $\overline{A}=C_{\mathbb{R}}(X)$，对于上述 $f\in C_{\mathbb{R}}(X)=\overline{A}$，任给 $\varepsilon>0$，$\exists\, g\in A$，使得

$$|f(z)-g(z)|<\varepsilon,\quad \forall\, z\in X. \tag{7.2.2}$$

于是，取 $\varepsilon=\dfrac{1}{2}$，结合(7.2.1)式与(7.2.2)式，有

$$-\frac{1}{2}<g(x)<\frac{1}{2},\quad \frac{1}{2}<g(y)<\frac{3}{2}. \tag{7.2.3}$$

从而 $g(x)\neq g(y)$．故 A 区分 X 的点．

充分性 若子代数 $A\subset C_{\mathbb{R}}(X)$区分 X 的点，证明 $\overline{A}=C_{\mathbb{R}}(X)$．分三步证明．

第一步，证明 $\overline{A}$ 在运算?,?：

$$(f\,?g)(x)=\max\{f(x),g(x)\},\quad (f\,?g)(x)=\min\{f(x),g(x)\}$$

之下封闭．

由于

$$\max\{f(x),g(x)\}=\frac{1}{2}\{f(x)+g(x)\}+\frac{1}{2}|f(x)-g(x)|,$$

$$\min\{f(x),g(x)\}=\frac{1}{2}\{f(x)+g(x)\}-\frac{1}{2}|f(x)-g(x)|,$$

为此，只需证明 $\overline{A}$ 对于取绝对值的映射 $\varphi:f\to|f|$是封闭的．

事实上，对于 $f\in\overline{A}\subset C_{\mathbb{R}}(X)\subset\mathfrak{B}(X)$，存在 $0<c<+\infty$，使得$|f|<c$．于是

$$|f|=\sqrt{c^2-(c^2-f^2)}=c\sqrt{1-\left(1-\frac{f^2}{c^2}\right)}=c\left\{1-\sum_{n=1}^{+\infty}\frac{(2n-3)!!}{(2n)!!}\left(1-\frac{f^2}{c^2}\right)^n\right\},$$

上式中的级数关于 $x\in X$ 一致收敛，并且因为 $f\in\overline{A}\subset C_{\mathbb{R}}(X)$，所以 f 连续，故每一项 $\left(1-\dfrac{f^2}{c^2}\right)^n$ 连续，从而级数和$|f|\in\overline{A}$．第一步得证．

第二步，证明 $\forall f\in C_{\mathbb{R}}(X)$ 及 $x,y\in X, x\neq y$，都存在 $g_{xy}\in A$，使得

$$g_{xy}(x)=f(x),\quad g_{xy}(y)=f(y).$$

事实上，对于 $x,y\in X, x\neq y$，由于子代数 A 区分 X，故存在 $g\in A$，使得 $g(x)\neq g(y)$，这等价于行列式 $\begin{vmatrix} g(x) & 1 \\ g(y) & 1 \end{vmatrix}\neq 0$，因此方程组

$$\begin{cases}\alpha g(x)+1\cdot\beta=f(x),\\ \alpha g(y)+1\cdot\beta=f(y)\end{cases}$$

在实数域 $\mathbb{R}$ 中有惟一的一组解 α,β. 令 $g_{xy}=\alpha g+\beta e$，这个 g_{xy} 就是满足要求的 $g_{xy}\in A$.

第三步，证明 $\overline{A}=C_{\mathbb{R}}(X)$，也只需证明 $C_{\mathbb{R}}(X)\subset\overline{A}$. 亦即，任取 $f\in C_{\mathbb{R}}(X)$，对于任给的 $\varepsilon>0$，存在 $g\in A$，使得 $\|f-g\|_X<\varepsilon$.

事实上，任给 $\varepsilon>0$，对于每个 $x\in X$，取一个 $y\in X, y\neq x$. 据第二步，存在 $g_{xy}\in A$，使得

$$g_{xy}(x)=f(x),\quad g_{xy}(y)=f(y).$$

由于 g_{xy} 的连续性，存在 $x\in X$ 的开邻域 $U_x\subset X$，使得 $-\varepsilon<g_{xy}(z)-f(z)<\varepsilon,\ z\in U_x$.

显然，$\bigcup\limits_{x\in X}U_x=X$，由 X 的紧致性，存在 $U_{x_1},\cdots,U_{x_n}$，使得 $\bigcup\limits_{j=1}^{n}U_{x_j}=X$. 于是，当 $z\in U_{x_j}$ 时，有 $g_{x_jy}(z)>f(z)-\varepsilon$. 令 $g_y=g_{x_1y}?\cdots?g_{x_ny}$，则

$$g_y(z)>f(z)-\varepsilon,\quad \forall z\in X=\bigcup_{j=1}^{n}U_{x_j}. \tag{7.2.4}$$

由第二步知，$g_y=g_{x_1y}?\cdots?g_{x_ny}\in A$. 于是，由 g_y 的连续性，存在 $y\in X$ 的开邻域 $V_y\subset X$，使得

$$-\varepsilon<g_y(z)-f(z)<\varepsilon,\quad z\in V_y.$$

显然，$\bigcup\limits_{y\in X}V_y=X$，由 X 的紧致性，存在 $V_{y_1},\cdots,V_{y_m}$，使得 $\bigcup\limits_{j=1}^{m}V_{y_j}=X$. 于是，当 $z\in V_{y_j}$ 时，有 $g_{y_j}(z)<f(z)+\varepsilon$. 令 $g=g_{y_1}?\cdots?g_{y_m}$，则

$$g(z)<f(z)+\varepsilon,\quad \forall z\in X=\bigcup_{j=1}^{m}V_{y_j}. \tag{7.2.5}$$

再由第二步，$g=g_{y_1}?\cdots?g_{y_m}\in A$.

联合(7.2.4)式与(7.2.5)式，可知对于任取的 $f\in C_{\mathbb{R}}(X)$，任给的 $\varepsilon>0$，存在 $g\in A$，使得

$$\|f-g\|_X=\max_{z\in X}|f(z)-g(z)|<\varepsilon.$$

定理得证.

定理 7.2.3(复数域上的 Stone-Weierstrass 定理) 设 X 为 T_2 型紧致拓扑空间，$\mathfrak{C}(X)$ 为 X 上复值连续函数所生成的 Banach 代数. 若 $A\subset\mathfrak{C}(X)$ 是 $\mathfrak{C}(X)$ 的含有 $\mathfrak{C}(X)$ 的乘法单位元 e 的子代数，则 $\overline{A}=\mathfrak{C}(X)$，当且仅当 A 区分 X 的点.

Stone-Weierstrass 定理有许多重要的应用. 读者可参看[10].

7.2.2 隐映射定理、逆映射定理

定理 7.2.4(微积分学中的隐函数存在定理) 假设

(1) 函数 $F(x,y)$ 在以点 (x_0,y_0) 为中心的某一邻域 $\mathfrak{D}=[x_0-\delta,x_0+\delta;y_0-\delta_1,y_0+\delta_1]$ 中有定义且连续;

(2) F'_x、F'_y 在 $\mathfrak{D}$ 中存在且连续;

(3) $F(x_0,y_0)=0$; $F'_y(x_0,y_0)\neq 0$.

则有

(1) 在 (x_0,y_0) 的某个邻域内,$F(x,y)=0$ 确定 y 为 x 的单值函数 $y=f(x)$;

(2) $y_0=f(x_0)$;

(3) $y=f(x)$ 在该邻域内连续可导,且 $\dfrac{\mathrm{d}y}{\mathrm{d}x}=-\dfrac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial y}}$.

定理 7.2.5(Banach 空间的隐映射定理) 设 X,Y,Z 为 Banach 空间,$A\subset X\times Y$ 为 $X\times Y$ 的开子集,若映射 $f:A\to Z$ 满足如下条件:

(1) f 在 A 上连续可导;

(2) 对于 $(x_0,y_0)\in A$,有 $f(x_0,y_0)=0$;

(3) $\partial_y f(x_0,y_0)$ 为 Y 到 Z 的线性同胚.

则存在 x_0 的开邻域 $U_0\subset X$,使得

(1) 对 U_0 的每个连通开子集 $U\subset U_0$,$x_0\in U$,存在惟一的连续映射 $u:U\to Y$,满足 $u(x_0)=y_0$,且对于每个 $x\in U$,有 $(x,u(x))\in A$ 与 $f(x,u(x))=0$;

(2) 上述的 u 在 U 中连续可导,且成立 $u'(x)=-(\partial_y f(x,u(x)))^{-1}\circ\partial_x f(x,u(x))$.

定理 7.2.6(Banach 空间中的逆映射定理) 设 X,Y 为 Banach 空间,$x_0\in X$,且 $V\subset X$ 为 x_0 的开邻域.若 $f:V\to Y$ 是连续可导映射,且 $f'(x_0)$ 为 X 到 Y 上的线性同胚,则存在 x_0 的开邻域 $U\subset V$,使得 $f|_U: U\to W\subset Y$ 为同胚映射,这里 W 为 $y_0=f(x_0)$ 的开邻域;并且若 $f:U\to W$ 在 U 内是 p 次连续可导的,则 $f:U\to f(U)\subset W$ 的逆映射 $g:f(U)\to U$ 在 $f(U)$ 中也是 p 次连续可导的.

注 Banach 空间中,映射可导的定义见[11]的第6章.

7.2.3 不动点原理

定理 7.2.7(压缩映象原理) 设 X,Y 为 Banach 空间,$U=B(0,a)\subset X$ 是 X 中以零元为心,以 $a>0$ 为半径的开球;$V=B(0,b)\subset Y$ 是 Y 中以零元为心,以 $b>0$ 为半径的开球;$u:U\times V\to Y$ 为连续映射.若 u 满足

(1) $\|u(x,y_1)-u(x,y_2)\|_{X\times Y}\leqslant k\|y_1-y_2\|_Y$,其中 $0<k<1$,$x\in U$,$y_1,y_2\in V$;

(2) $\|u(x,0)\|_{X\times Y}<\beta(1-k)$,其中 $x\in U$,
则存在惟一的映射 $f:U\to V$,满足 $f(x)=u(x,f(x))$.

高等数学中的常微分方程初始值问题的解的存在惟一性定理(见[11]的第 5 章)就是本定理的特殊情形.

定理 7.2.8(Brouwer 不动点原理) 设 $B(0,1)=\{x\in\mathbb{R}^n:\|x\|_{\mathbb{R}^n}\leqslant 1\}$ 为$\mathbb{R}^n$ 中的单位闭球,若此球上的映射 $f:B(0,1)\to B(0,1)$ 为连续映射,则存在 $x_0\in B(0,1)$,使得 $f(x_0)=x_0$.

Schauder 不动点原理的两种常用的形式:

定理 7.2.9(有界闭凸集与紧映射的不动点原理) 设 X 为 Banach 空间,$M\subset X$ 是 X 中的有界闭凸集,若 $f:M\to M$ 为紧映射,则存在 $x_0\in M$,使得 $f(x_0)=x_0$.

注 1 Banach 空间 X 中的凸集 M 是指:对于任意两点 $x_1,x_2\in M$,若连接两点的“直线”

$$L=\{x\in X:x=\alpha x_1+(1-\alpha)x_2,0\leqslant\alpha\leqslant 1\}$$

全部包含在 M 中,即 $L\subset M$.

注 2 Banach 空间 X 中的子集 $M\subseteq X$ 上的紧映射 $f:M\to M$ 是指:对于任意序列 $\{x_n\}\subset M$,映射序列$\{f(x_n)\}\subset M$ 都有收敛到 M 中的元的子序列.

定理 7.2.10(紧凸集与连续映射的不动点原理) 设 X 为 Banach 空间,$M\subset X$ 为 X 的紧凸子集,若 $f:M\to M$ 为连续映射,则存在 $x_0\in M$,使得 $f(x_0)=x_0$.

7.3 局部紧群上的 Haar 积分

Haar 测度是 1933 年由德国数学家 A. Haar 首先建立的,是 Lebesgue 测度的一种推广,从而建立相应的 Haar 积分. 这类测度(积分)具有平移不变性,概括了一类具有刚性运动物体的“体积”不变的度量,因而成为数学科学、物理学以及自然科学中多学科的重要工具.

1. 局部紧群,紧群

定义 7.3.1(拓扑群) 设集合 G 的运算结构是一个群,拓扑结构是一个拓扑空间. 若 $G\times G$ 到 G 的映射$(x,y)\to xy$ 与 G 到 G 的映射 $x\to x^{-1}$是连续的,则称 G 是一个**拓扑群**.

拓扑群的例子很多,例如,实数集合$\mathbb{R}$ 在通常的加法+与通常拓扑 τ 之下构成一个拓扑群. 正实数集合$\mathbb{R}^+=(0,+\infty)$在通常乘法×与通常拓扑 τ 之下构成一个拓扑群等.

定义 7.3.2(局部紧致拓扑群,紧致拓扑群) 设 G 是一个拓扑群,若其拓扑结构 τ 使 G 成为局部紧致拓扑空间,则称 G 为一个**局部紧致拓扑群**,简称**局部紧群**;若 τ 使 G 成为紧致拓扑空间,则称 G 为**紧致拓扑群**,简称**紧群**.

2. 局部紧群上的Haar积分

定义 7.3.3(**G 上的正线性泛函、Haar 积分**) 设 G 是一个 T_2 型局部紧群，$\mathfrak{C}(G)$ 是 G 上具有紧致支集的复值连续函数的全体，我们称 $\mathfrak{C}(G)$ 上的线性泛函 I 为 G 上的**左 Haar 积分**，若

(1) $f\in\mathfrak{C}(G)$，$f>0$ 蕴含 $I(f)>0$；

(2) $f\in\mathfrak{C}(G)$，$f\neq 0$，$\forall a\in G$ 蕴含 $I({}_af)=I(f)$，

其中${}_af(x)=f(ax)$. 同样，可定义右 Haar 积分.

当线性泛函 I 既是左 Haar 积分，又是右 Haar 积分时，就称其为 G 上的 Haar 积分. 显然，当 G 是交换群时，其上的左、右 Haar 积分就是 Haar 积分.

定理 7.3.1(**局部紧群上 Haar 积分的存在性**) 设 G 是 T_2 型局部紧群，则存在 G 上的左(右)Haar 积分；若不计常数因子，此左(右)Haar 积分是惟一的. 当 G 是交换群时，其上的 Haar 积分，若不计常数因子，是惟一的.

设 G 为 T_2 型局部紧群，$f:G\to\mathbb{C}$ 是 G 到复数域$\mathbb{C}$ 的复值函数，f 的 Haar 积分(或左、或右)记为

$$\int_G f(x)\mathrm{d}x. \tag{7.3.1}$$

3. 例子

例 7.3.1 设$(\mathbb{R}^n,+)$为欧氏空间$\mathbb{R}^n$ 在加法$+$运算下所构成的 T_2 型局部紧 Abel 群，其上的 Haar 积分就是通常的 n 重 Lebesgue 积分 $\int_{\mathbb{R}^n}f(x)\mathrm{d}m$.

例 7.3.2 设$(G,\cdot)=(\{x_1,\cdots,x_r\},\cdot)$是 r 阶有限 Abel 群，则线性泛函 $I(f)=\frac{1}{r}\sum\limits_{j=1}^{r}f(x_j)$是 G 上的(左与右)Haar 积分.

例 7.3.3 设$(\mathbb{R}^+,\times)=((0,+\infty),\times)$为正实轴$\mathbb{R}^+$ 在乘法$\times$运算下所构成的 T_2 型局部紧 Abel 群，则

$$\int_{\mathbb{R}^+}f(x)\frac{\mathrm{d}x}{x}$$

是$\mathbb{R}^+$ 上的 Haar 积分.

4. 局部紧群上的Haar测度

定义 7.3.4(**Borel 集类**) 设 G 是一个 T_2 型局部紧群，G 的开子集、闭子集以及由开、闭子集的并集与交集所成的集合的全体，称为 G 上的 **Borel 集类**，记为 $\mathfrak{R}G)$. $\mathfrak{R}G)$中的集(或是 G 的开集，或是 G 的闭集，或是开集的任意交，或是闭集的任意并)称为 **Borel 集.**

定义 7.3.5(Haar 测度) 对于 T_2 型局部紧群 G,若存在 Borel 集类 $\mathfrak{B}(G)$上的一个非负实值集函数 μ,使得 $\forall E\in\mathfrak{B}(G)$,满足

(1) $\mu(\varnothing)=0$;

(2) 对于 $E_j\in\mathfrak{B}(G)$,$j=1,2,\cdots$,当 $j\neq k$ 时 $E_j\cap E_k=\varnothing$,有

$$\mu\left(\bigcup_{j=1}^{+\infty}E_j\right)=\sum_{j=1}^{+\infty}\mu(E_j),\tag{7.3.2}$$

则称 μ 为 $\mathfrak{B}(G)$上的 **Borel 测度**.

若 Borel 测度 μ 满足:对于任意 $x\in G$ 与 $E\in\mathfrak{B}(G)$,有

$$\mu(xE)=\mu(E)\quad(\text{或 }\mu(Ex)=\mu(E)),$$

则称 μ 为 $\mathfrak{B}(G)$上的**左平移不变测度**(或**右平移不变测度**),也称为左 Haar 测度(或右 Haar 测度). 若 μ 既是 $\mathfrak{B}(G)$上的左 Haar 测度,又是 $\mathfrak{B}(G)$上的右 Haar 测度,则称 μ 为 $\mathfrak{B}(G)$上的 Haar 测度.

5. 局部紧群上的 Haar 测度与 Haar 积分的关系

对于 T_2 型局部紧群 G 上的一个 Haar 积分(7.3.1)式,若集合 $E\in\mathfrak{B}(G)$是 Borel 可测集,且 $\mu(E)$是 E 的 Borel 测度,则对于 E 的"特征函数"

$$\chi_E(x)=\begin{cases}1, & x\in E,\\ 0, & x\notin E;\end{cases}$$

可定义 E 的一个 Haar 测度

$$\nu(E)=\int_G\chi_E(x)\mathrm{d}x,$$

此测度与 $\mu(E)$至多相差一个常数因子 c,$\mu(E)=c\nu(E)$,$\forall E\in\mathfrak{B}(G)$.

反之,对于 T_2 型局部紧群 G 上的一个 Haar 测度 μ,可定义一个 Haar 积分

$$I(f)=\int_G f(x)\mathrm{d}\mu(x),$$

其中 $\mathrm{d}\mu$ 是关于测度 μ 的积分微元.

6. Haar 积分的性质

定理 7.3.2(局部紧群上 Haar 积分的性质) 设 G 是 T_2 型局部紧群,若 μ 是 G 上的 Haar 测度,则

(1) 对于 G 的非空开集 $O\subset G$,有 $\mu(O)>0$;

(2) 对于 G 上的正的具紧致支集的连续函数 $f>0$,有 $\int_G f\mathrm{d}\mu>0$;

(3) $\mu(G)<+\infty$,当且仅当 G 为紧群.

定义 7.3.6(Haar 测度的模函数) 设 G 是一个 T_2 型局部紧群,I 是 G 上的 Haar 积分,对于 $f\in\mathfrak{C}(G)$,$f>0$ 与 $x\in G$,令 $\Delta(x)=\dfrac{I(f_{x^{-1}})}{I(f)}$,称 $\Delta(x)$为 G 的**模函数**.

定理 7.3.3(模函数的性质) 设 G 是 T_2 型局部紧群，若 $\Delta(x)=\dfrac{I(f_{x^{-1}})}{I(f)}$ 是 G 的模函数，则

(1) $\Delta(x)$ 仅依赖于 x，与 f，I 无关；

(2) $\Delta(x)$ 是 G 上的正值连续函数，且满足 $\Delta(x\cdot y)=\Delta(x)\Delta(y)$；

(3) 每个紧群 G 是么模的，亦即 $\Delta(x)=1$.

习题 7

1. 在点(x,u)处，介质的折射率为 $n(x,u)$，光通过两点间需时最小的路径就是实际的光线，这就是 Fermat 原理. 试将它改述成变分问题(光的速度与 $n(x,u)$ 成正比).

2. 推导泛函 $J(u)=\int_{a_0}^{a_1}F(x,u_1,\cdots,u_k,u'_1,\cdots,u'_k)\mathrm{d}x$ 的 Euler 方程(7.1.18).

3. 推导泛函 $J(u)=\int_{a_0}^{a_1}F(x,u,u',\cdots,u^{(k)})\mathrm{d}x$ 的 Euler 方程(7.1.19).

4. 推导泛函 $J(u)=\iint\limits_G F(x,y,u,u_x,u_y)\mathrm{d}x\mathrm{d}y$ 的 Euler 方程(7.1.20).

5. 在捷线问题中，求极小值 $J(\tilde{u})$，请利用对应于 P_1 的角 θ_1 来表示.

6. 试求 $J(u)=\int u^k(u^2+(u')^2)^{\frac{1}{2}}\mathrm{d}x, u>0$ 的极小函数与极小值.

7. 试证：形如 $J(u)=\int u^k g(u')\mathrm{d}x$ 的泛函的极值函数具有 $u=c\varphi\left(\dfrac{x-\xi}{c}\right)$ 的形式，其中 c,ξ 为任意常数.

8. 考虑泛函 $J(u)=\int_{a_0}^{a_1}[p(x)u^2+2q(x)uu'+r(x)u'^2+2f(x)u+2g(x)u']\mathrm{d}x$ 的极值问题该如何解决？

9. 考虑泛函 $J(u)=\iint\limits_\Omega[u_x^2+u_y^2+2f(x,y)u]\mathrm{d}x\mathrm{d}y$ 的极值问题该如何解决？

10. 考虑泛函 $J(u)=\iint\limits_\Omega[u_x^2+u_y^2+2f(u)]\mathrm{d}x\mathrm{d}y+\int_{\partial\Omega}[\rho(s)u^2+2\sigma(s)u]\mathrm{d}s$ 的极值问题该如何解决？

11. 解固有值问题 $J(u)=\int_{a_0}^{a_1}[p(x)u^2+2q(x)uu'+r(x)(u')^2]\mathrm{d}x$ 在附加条件 $K(u)=\int_{a_0}^{a_1}k(x)u^2\mathrm{d}x=1$ 之下的最小值问题(等周问题).

12. 导出 $J(u)=\iint\limits_\Omega(u_x^2-u_y^2)\mathrm{d}x\mathrm{d}y$ 的 Euler 方程.

参 考 文 献

[1] 张禾瑞.近世代数基础[M].北京：人民教育出版社，1978.

[2] 迪厄多内J.现代分析基础：第一卷[M].郭瑞芝，苏维宜，译.北京：科学出版社，1983.

[3] 陈省身，陈维桓.微分几何讲义[M].北京：北京大学出版社，1984.

[4] 巴策P L，等.Fourier分析与逼近论(第一卷，上册)[M].郑维行，苏维宜，任福贤，等，译.北京：高等教育出版社，1985.

[5] 齐民友.线性偏微分算子引论(上)[M].北京：科学出版社，1986.

[6] 郑维行，王声望.实变函数与泛函分析概要(一)、(二)[M].北京：高等教育出版社，1992.

[7] 吴时敏.广义相对论教程[M].北京：北京师范大学出版社，1998.

[8] 齐民友，吴同方.广义函数与数学物理方程[M].北京：高等教育出版社，1999.

[9] 陈维桓.流形上的微积分[M].北京：高等教育出版社，2000.

[10] 苏维宜.近代分析引论[M].北京：北京大学出版社，2000.

[11] 仇庆久.高等数学(上、下册)[M].北京：高等教育出版社，2003.

[12] 熊金城.点集拓扑讲义[M].北京：高等教育出版社，2003.

[13] 龚昇.线性代数五讲[M].北京：科学出版社，2005.

[14] 徐森林，等.点集拓扑学[M].北京：高等教育出版社，2007.

[15] 曾谨言.量子力学教程[M].北京：科学出版社，2008.

[16] 徐森林，等.近代微分几何[M].合肥：中国科学技术大学出版社，2009.

[17] 徐森林，等.实变函数论[M].北京：清华大学出版社，2009.

[18] 苏维宜.局部域上的调和分析与分形分析[M].北京：科学出版社，2011.